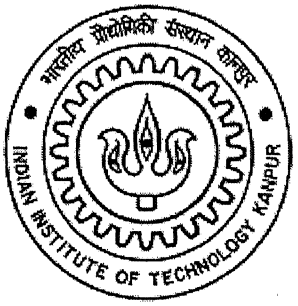
PARAMETER ESTIMATION AND MODEL SELECTION IN SINGLE AND MULTI RESPONSE SYSTEMS

A Thesis submitted
in partial fulfilment of the requirements
for the Degree of

Master of Technology

by

Dinesh Kumar Bolisetty



to the

DEPARTMENT OF CHEMICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY KANPUR

June 2004

28 JUL 2004 / CHE
पुरुषोत्तम कालीनाथ केलकर पुस्तकालय
भारतीय प्रौद्योगिकी संस्थान कानपुर
प्राप्ति क्र० A... 148448
TH
CHE/2004/M
B638P

Dedicated

to

My Venerable and

Beloved

Family Members

POST GRADUATE OFFICE
THESIS SUBMITTED
ON 11/6/04
I.I.T., KANPUR 16

CERTIFICATE

This is to certify that the present work entitled 'PARAMETER ESTIMATION AND MODEL SELECTION IN SINGLE AND MULTIRESPONSE SYSTEMS' has been carried out by DINESH KUMAR BOLISETTY under our supervision and this work has not been submitted elsewhere for a degree.

Dr. D. Kundu

Professor
Department of Mathematics
Indian Institute of Technology
Kanpur-208016, India.

Dr. M. Someswara Rao

Professor
Department of Chemical Engineering
Indian Institute of Technology
Kanpur-208016, India.

JUNE, 2004

ACKNOWLEDGMENT

I would like to express my deep sense of gratitude towards my Thesis Supervisor Dr. M. S. Rao for his valuable guidance, support and constant encouragement throughout my thesis work. I would take this opportunity to express my gratefulness and dedications to my co-supervisor Dr. D. Kundu for his constructive criticism, untiring help and invaluable advice given to me during the crucial period of my thesis.

I thank to all chemical engineering faculty members for providing such a nice course and research oriented program, which opened new horizons of success for me.

I would also like to thank Mr. R. A. Mishra for their help during my thesis work. I am fortunate to have such a nice and sincere labmates like Sushil and seniors Ravi, Negi, Rajeev and Vishal with whom I spent a considerable part of my life.

I am grateful to God for giving me such a large, amicable and co-operating friend circle that includes many persons thus to name one or two would be very unfair with others. However to name some of few I am very thankful to Mallesh, Pavan, Ajith, Nages, Sameer, Simha, Pavan, Sudhakar, Krishna, Nagesh, Verma, Gullabani, Somesh and Prince for their great affection, care and timely criticism.

Finally, to my parents, my beloved sister, goes my eternal gratitude for their constant love and support.

Dinesh Kumar Bolisetty

CONTENTS

LIST OF FIGURES		vi
LIST OF TABLES		vii
NOMENCLATURE		viii
ABSTRACT		xi

CHAPTER		PAGE
1	**INTRODUCTION**	1
1.1	Uses of Mathematical Modeling	2
1.2	Process of Model Building	2
1.2.1	Parameter estimation	3
1.2.2	Model Selection	5
1.3	Single and Multi response Models	6
1.4	Objective of the Present Investigation	7
1.5	Thesis Organization	7
2	**LITERATURE REVIEW**	8
2.1	Review of the Previous Work : Point Estimations	8
2.2	Review of the Previous Work : Interval Estimations	10
2.3	Review of the Previous Work in Model Selection	11
2.4	Summary	12
3	**METHODOLOGY**	13
3.1	Single Response Models	13
3.1.1	Simulation of the Data	14
3.1.2	Parameter Estimation	14
3.1.3	Model Selection	14
3.2	Multi Response Models	15
3.2.1	Simulation of the Data	16
3.2.2	Parameter Estimation	16
3.2.3	Model Selection	17
3.3	GNLM Method for Parameter Estimation	18
3.4	Asymptotic Properties of Non-Linear Least Squares Estimation	24
3.5	Bootstrapping Methods for Parameter Estimation	26
3.5.1	General Bootstrap Methodology	26
3.5.2	Bootstrap Percentile Confidence Interval	28

3.6 Cross Validation Method for Model Selection 29
3.7 Estimation of Average Estimates and MSE of the Parameters 30

4 RESULTS AND DISCUSSIONS 31

4.1 Single Response System 31
4.1.1 Simulation of the Data 31
4.1.2 Parameter Estimation 32
4.1.3 Estimation of Average Estimates and MSE of the Parameters 33
4.1.4 Model Selection 37
4.2 Multi Response System 37
4.2.1 Simulation of the Data 37
4.2.2 Parameter Estimation 38
4.2.3 Estimation of Average Estimates and MSE of the Parameters 39
4.2.4 Model Selection 58
4.3 Application to Real Example 58
4.3.1 Single Response System 58
4.3.2 Parameter Estimation 59
4.3.3 Model Selection 60

5 CONCLUSIONS & RECOMMENDATIONS 69

5.1 Conclusions 69
5.2 Recommendations 70

BIBLIOGRAPHY 71

APPENDIX I 75

APPENDIX II 77

APPENDIX III 79

APPENDIX IV 82

APPENDIX V 86

APPENDIX VI 87

APPENDIX VII 88

LIST OF FIGURES

		PAGE
1.1	Modeling Strategy	3
3.1	Example for a Multi Response Model	15
3.2	GNLM Method Flow Diagram	23
3.3	Flow Diagram for Bootstrap Replications	27
4.1	Histogram & Q-Q Plot for Parameter ξ_1	35
4.2	Histogram & Q-Q Plot for Parameter ξ_2	36
4.3	Histogram & Q-Q Plot for Parameter K_1 (n=10)	46
4.4	Histogram & Q-Q Plot for Parameter K_2 (n=10)	47
4.5	Histogram & Q-Q Plot for Parameter K_3 (n=10)	48
4.6	Histogram & Q-Q Plot for Parameter K_4 (n=10)	49
4.7	Histogram & Q-Q Plot for Parameter K_1 (n=15)	50
4.8	Histogram & Q-Q Plot for Parameter K_2 (n=15)	51
4.9	Histogram & Q-Q Plot for Parameter K_3 (n=15)	52
4.10	Histogram & Q-Q Plot for Parameter K_4 (n=15)	53
4.11	Histogram & Q-Q Plot for Parameter K_1 (n=19)	54
4.12	Histogram & Q-Q Plot for Parameter K_2 (n=19)	55
4.13	Histogram & Q-Q Plot for Parameter K_3 (n=19)	56
4.14	Histogram & Q-Q Plot for Parameter K_4 (n=19)	57
4.15	Histogram & Q-Q Plot for Parameter k_0	61
4.16	Histogram & Q-Q Plot for Parameter k_1	62
4.17	Histogram & Q-Q Plot for Parameter k_2	63
4.18	Histogram & Q-Q Plot for Parameter k_3	64
4.19	Histogram & Q-Q Plot for Parameter k_0	65
4.20	Histogram & Q-Q Plot for Parameter k_1	66
4.21	Histogram & Q-Q Plot for Parameter k_2	67
4.22	Histogram & Q-Q Plot for Parameter k_3	68

LIST OF TABLES

		PAGE
4.1	Single Response Data	31
4.2	The Simulated Response Data with different values of Variances	32
4.3	Estimated values of the Parameters	33
4.4	Simulation Results	34
4.5	Cross Validation Results	37
4.6	Multi Response Data	38
4.7	Estimated values of the Parameters	39
4.8	Mean Parameter Estimates	40
4.9	Estimates of Mean Squared Errors	41
4.10	Estimates of Mean Length of Asymptotic Confidence Intervals	42
4.11	Estimates of Mean Length of Bootstrap Confidence Intervals	43
4.12	Coverage Probabilities of Asymptotic Confidence Intervals	44
4.13	Coverage Probabilities of Bootstrap Confidence Intervals	45
4.14	Cross Validation Results	58
4.15	Comparison of the Estimated Parameters	59
4.16	Simulation Results	60

NOMENCLATURE

Notations

A'	transpose of A
A,B,C,D,E,F	chemical species A,B,C,D,E,F respectively involved in the reaction
C_i	concentration of the i^{th} species
CV	cross validation index
D_e	determinant criterion
$E(\varepsilon)$	expectation of ε
f	functional relationship
$\dot{F}(\theta)$	matrix of first derivative of $f(\theta)$ with respect to θ
$g(\theta)$	first derivative of residual sum of squares with respect to θ
$H(\theta)$	Hessian matrix (second derivative of residual sum of squares with respect to θ)
I	Identity matrix
J	Jacobean matrix
k_i	reaction rate constant of i^{th} forward reaction
k_{-i}	reaction rate constant of i^{th} backward reaction
$(K_j)_i$	adsorption constant of 'j' chemical species in the i^{th} reaction
n	number of experiments
N	Normal distribution
O.F.	objective function
Q	Reciprocal of covariance matrix
$-r_i$	rate of depletion of 'i' chemical species
r_i	rate of formation of 'i' chemical species
S	residual sum of squares function
SOS	residual sum of squares
$V(\varepsilon)$	variance of ε
x	independent variable
X	vector X

y	dependent variable
y_i	generated value of rate for chemical species 'i'
Y	vector Y
Z	matrix

Greek Symbols

β_i	i^{th} parameter in the model
δ	step length in Gauss Newton Method
γ	tolerance value
ε	error
λ	constant used in
η	true value Levenberg Marquardt Method
σ_{ij}	i,j^{th} element of variance-covariance matrix
σ^{ij}	i,j^{th} element of the inverse of variance-covariance matrix
σ^2	variance
Σ	variance-covariance matrix
θ_i	i^{th} parameters in a model
ξ_I	i^{th} variable in a model
μ	mean
φ	SOS in Levenberg Marquardt Method

Superscripts

(i)	i^{th} response
T	transpose
(u)	u^{th} experiment as in the Eqn. (1.12)
*	bootstrap replications in variables
b	bootstrap repliactes

Subscripts

i	number of row
j	number of column
B	bootstrap replications

At top

^	least square estimate
^*b	bootstrap estimates

At bottom

~	vector

ABSTRACT

Model building is a very important process in any scientific investigation. The researcher requires a mechanistic model, which is capable of describing the process completely. The final form of the mechanistic model may be obtained by following several steps carefully. The first and the foremost step, viz., the estimation of parameters is well recognized and the next is model selection. It is a process to obtain the unknown parameters in the model with the greatest amount of precision possible. One way to estimate the parameters for a large variety of models is to use the least squares technique, which involves minimizing the residual sum of squares of the differences between the experimental and the model values. In large majority of real systems, the investigator frequently comes across situations where several models have been postulated for the same system. This necessitates the discrimination among the competing models and selection of the best model.

In the present investigation the Asymptotic theory and Bootstrapping methods are employed for the parameter estimation process in the single response as well the multi response systems. Initially simulated data are used for the estimation of the parameters and the efficiency of the above mentioned methods are tested. The model selection is also carried out by using Cross Validation procedure. Using a real life data the Asymptotic theory and Bootstrapping methods are again used for parameter estimation in single response models and the estimates of the parameters are compared with the values given by the investigators.

CHAPTER 1

INTRODUCTION

In the study of a system one often wishes to know the behavior of the process. The best way to analyze a process is to put it into the mathematical equations, which can suitably describe the underlying phenomenon. This process is known as mathematical model building, which is a widely used tool and not untouched by any stream. In model building there is a dependent variable y_i (whose true value is η_i) related to a set of independent variables $\underset{\sim}{x} = (x_1, x_2, x_3, \ldots)$ and a vector of parameters $\underset{\sim}{\beta} = (\beta_1, \beta_2, \beta_3 \ldots)'$ of the process under study such that the equation shows the true nature of the process. Any phenomena can be represented mathematically by

$$\eta = f(\underset{\sim}{x}, \underset{\sim}{\beta}). \tag{1.1}$$

where η is the observed response, $\underset{\sim}{x}$ is the vector of independent variables and $\underset{\sim}{\beta}$ is the vector of parameters in the equation and η is related to $\underset{\sim}{x}$ and $\underset{\sim}{\beta}$ by function f.

This is known as the true mathematical model. In practice, the true mathematical model of a process is never precisely known. In many situations, the mechanism of a process is not completely understood and a very close mechanistic mathematical model representation is available. On the other hand, many situations occur where the researcher often wishes to obtain an approximate idea of the behavior of the response over some region of interest. When the mechanism is complex and its knowledge involves great effort, it may be wasteful to attempt to determine a mechanistic modely. In such cases the empirical approach of model building is used which does not need the knowledge of the mechanism. The response or dependent variable may safely be taken as a polynomial $f(\underset{\sim}{x}, \underset{\sim}{\beta})$ in the smooth region of interest

$$y = f(\underset{\sim}{x}, \underset{\sim}{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \beta_{12} x_1 x_2 + \ldots$$

$$+\beta_{11} x_1^2 + \ldots + \beta_{kk} x_k^2 + \varepsilon \tag{1.2}$$

where $\underset{\sim}{\beta}$ is a vector of parameters, $\underset{\sim}{x}$ is the vector of independent variables, y is the observed response and ε is the vector of error values.

The experiments are conducted in a region of interest and an adequate polynomial fitted by the least square method. It is not suggested that an empirical approach is used in cases where the objective is to know the mechanism, when this can be done simply by mechanistic modeling.

1.1 Uses of Mathematical Modeling

The mathematical modeling process is generally used in each and every scientific investigation. It is a very useful tool to compress large amount of information, providing new insights into the process and suggesting ways for the future development of the system.

Often there exists a functional relationship, which is too complicated to grasp or to describe in simple terms. In this case, the experimenter may wish to approximate this functional relationship by some simple mathematical function, such as a polynomial, which contains the appropriate variables and which graduates or approximates to the true function over some limited ranges of variables involved. By examining such a graduating function, the investigator may be able to learn more about the underlying true relationship and to understand the effects produced by changes in certain important variables.

It becomes possible by model building to predict responses and optimize the system. It is very important in industrial processes when it is very important to have scientific knowledge of what is actually happening. The design, optimization, control, maintenance etc. all steps become simpler by analyzing mathematical equation.

1.2 Process of Model Building

The process of modeling is an important operation and to be carried out carefully in stages. Box and Hunter [7] and Kanodia [30] also represented these steps by the following figure.

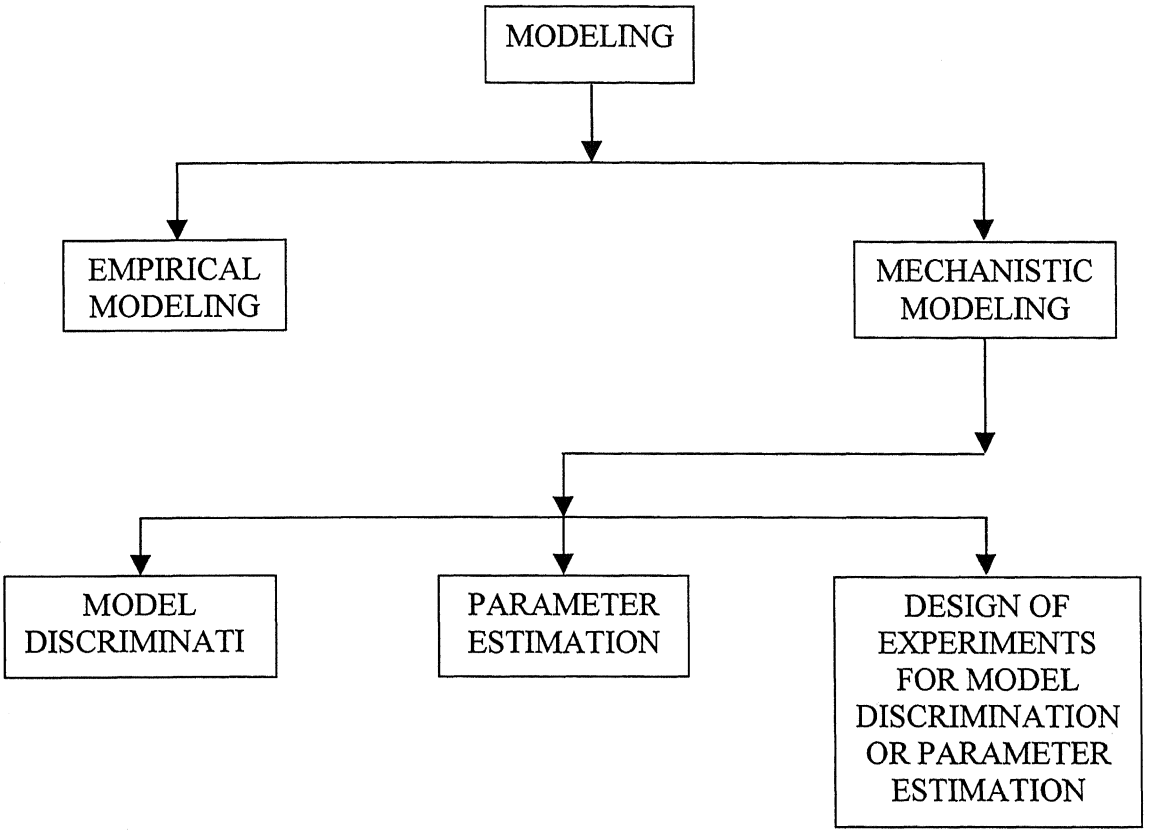


Fig 1.1 Modeling Strategy

The work of model building can be categorized in the two areas. (1) Parameter estimation and (2) Model specification.

1.2.1 Parameter estimation

The experimenter conducts some preliminary experiments and collects some data from these experiments. If there happens to be a single model considered appropriate for describing the process of the system, the only job of the experimenter is to calculate the precise estimates of the parameters i.e. the unknowns appearing in the model equation.

A chemical engineer deals regularly with the experimental data in studies of chemical reactions, mixtures and processes. If the rate-determining step is known a mechanistic model can be derived from this information. The rate equation involves several unknown parameters like reaction rate constants and adsorption equilibrium constants depending upon the rate equation.

The method of least squares (applied regression analysis) is well recognized for the estimation of parameters. The unknown parameters are estimated in the light of experimental data. Suppose we have n observations ($\underset{\sim}{x}_i$, y_i), i = 1, 2,...n, from a fixed regression model with a known functional relationship f. Then

$$y_i = f(\underset{\sim}{x}_i, \underset{\sim}{\beta}^*) + \varepsilon_i, \tag{1.3}$$

where it is assumed that E (ε_i) = 0 and $v(\underset{\sim}{\varepsilon}) = I\sigma^2$. The least square estimates of β can be obtained by minimizing the error sum of squares, $S(\beta)$ given by

$$S(\beta) = \sum_{i=1}^{n} \left[y_i - f(\underset{\sim}{x}_i, \underset{\sim}{\beta}) \right]^2 . \tag{1.4}$$

By minimizing this residual sum of squares with respect to $\underset{\sim}{\beta}$ with the aid of one of several optimization techniques, the experimenter can estimate the least squares estimate of the parameter β. The geometrical interpretation of theory of least squares estimation can be better understood by representing the equation in vector notation.

$$\underset{\sim}{y} = f(\underset{\sim}{X}, \underset{\sim}{\beta}) + \underset{\sim}{\varepsilon} \tag{1.5}$$

where $\underset{\sim}{y}$ and $\underset{\sim}{\varepsilon}$ are (n x 1) vectors and $\underset{\sim}{X}$ is (n x k) design matrix and f is linear function.

This method minimizes $\sum_i \varepsilon_i^2$ with respect to $\underset{\sim}{\beta}$, i.e. to minimize

$$S = (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})^T (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta}) . \tag{1.6}$$

Or

$$S = \left\| \underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta} \right\|. \tag{1.7}$$

$\left\| \underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta} \right\|^2$ (the square of the length of $\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta}$) will be minimum for $\underset{\sim}{X}\underset{\sim}{\beta} = \underset{\sim}{X}\underset{\sim}{\hat{\beta}}$, when $(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})$ will be orthogonal to X space. In mathematical representation

$$\underset{\sim}{X}^T (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta}) = 0 \tag{1.8}$$

$$\underset{\sim}{X}^T\underset{\sim}{y} - \underset{\sim}{X}^T\underset{\sim}{X}\underset{\sim}{\beta} = 0$$

$$\underset{\sim}{\hat{\beta}} = (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{y} \qquad (1.9)$$

This will give the least square estimates of $\underset{\sim}{\beta}$.

1.2.2 Model Specification

In the previous section it was assumed that the correct form of the model was known. Frequently in many situations involving chemical kinetics the underlying physical theory suggests several rival models are capable of explaining the mechanism of the system. In such situations, a model can be actually put to use only if a single model is picked out of the given lot and fitted precisely to the available data. This problem is commonly referred as model-discrimination problem.

In order to reduce a large number of models to a manageable group the process generally used is known as model screening. Two widely used methods are lack of fit and residual plots. Linear models showing significant lack of fit can be eliminated simply by F-test on residual sum of squares. The other method, which uses residual plots can be employed to indicate trends peculiar to the model.

The methods used for discrimination criteria can be categorized as Bayesian methods and the classical methods. The Bayesian method is based upon the Bayes theorem. A value is assigned to the prior probability of each of the model and then expected likelihood of each model is calculated and on comparing these values the models are discriminated. The classical methods use likelihood ratio approach. The different information theoretic criteria with proper choice of penalty function are widely used to choose the correct model.

Sometimes the experimenter can not discriminate between the two equally good rival models from the existing data. In this case the experimenter is required to carefully design the experiments to distinguish among the rival models and the design of experiments should be in the region of interest, where the discrimination among the rival models is maximum.

It may happen after screening and discrimination analysis that a model is chosen that shows specific weaknesses under the certain special conditions. To correct

for these weaknesses the model should be modified until a form of the model is found that is adequate over the operable conditions for system under study.

1.3 Single and Multiresponse Models

Generally the true value of the response is given by the Equation (1.1). Suppose y is the observed response, $\underset{\sim}{x}$ is the vector of independent variables and $\underset{\sim}{\beta}$ is the vector of parameters in the equation and y is related to $\underset{\sim}{x}$ and $\underset{\sim}{\beta}$ by function f, then

$$y = f(\underset{\sim}{x}, \underset{\sim}{\beta}) + \varepsilon, \tag{1.10}$$

where ε is the error associated with the measurement. If for a particular system only one dependent variable is observed, the system is known as single response system.

However, situations are possible when the experimenter observes more than one response in a system. As in chemical reaction engineering, kinetic investigation frequently involves with the models where more than one rate equation is possible for a reaction. In such situations the experimenter has to observe several responses instead of one and this multiresponse model can be represented by

$$\underset{\sim}{Y} = \underset{\sim}{f}(\underset{\sim}{X}, \underset{\sim}{\beta}) + \underset{\sim}{\varepsilon}, \tag{1.11}$$

where the i-th response $(1 \le i \le r)$ for the u-th experiment $(1 \le u \le n)$ may be denoted by

$$y_u^{(i)} = \eta_u^{(i)}(\underset{\sim}{X}_u, \underset{\sim}{\beta}) + \varepsilon_u^{(i)}. \tag{1.12}$$

It is assumed by Rao and Iyengar [36] that the errors are such that

$$\begin{aligned} &E(\varepsilon_u^{(i)}) = 0 \text{ for all i, u} \\ &E(\varepsilon_u^{(i)} \varepsilon_v^{(j)}) = 0 \text{ for all i, j, } u \neq v \\ &E(\varepsilon_u^{(i)} \varepsilon_u^{(j)}) = \sigma_{ij} \text{ for i, j, u} \end{aligned} \tag{1.13}$$

The vector $\underset{\sim}{y}_u = (y_u^{(1)}, y_u^{(2)} \ldots y_u^{(r)})'$ of r responses for the u^{th} experiments has a symmetric covariance matrix $\sum$ given by

$$\sum = \left((\sigma_{ij}) \right). \qquad i = 1, 2, \ldots r \ \& \ j = 1, 2, \ldots r. \tag{1.14}$$

1.4 Objective of the present investigation

The objective of the present investigation is to devise a reliable method for the estimation of unbiased parameters in both single response and multi-response models. Initially simulated data are used for estimation of parameters from a known model for both single response and multiresponse models and the differences between Point estimation methods and Interval estimation methods for parameter estimation has been tested. Then the method is extended to real life problems discussed by carr [11] and compare with the values of parameters obtained by the investigators.

1.5 Thesis Organization

This thesis consists five chapters of which the current chapter is the first. The next chapter, Chapter 2, presents a brief literature review related to the present study. The methods used in the present study are explained in Chapter 3. In Chapter 4, the results and discussion of present work is given. Conclusions and recommendations for future work are reported in Chapter 5. Finally the references used throughout the thesis are given in alphabetical order.

CHAPTER 2

LITERATURE REVIEW

In the past few years most of the studies were made on the model discrimination and design of experiments than parameter estimation. But the parameter estimation is also very important in model discrimination and design of experiments. Estimation in algebraic equations, linear in parameters, is well known, and elementary computer packages contain all the associated statistical tests.

Though attempts have been made to find out the reaction kinetics for different types of reactions, selection of objective function for nonlinear system has always been a matter of contention. Various methods for estimation of parameters in case of single and multi response systems are available in literature, but the choice of using objective function for nonlinear regression is quite interesting. Also there are number of optimization methods are available for the estimation of parameters by minimizing objective function. In recent years the breakthroughs in statistics is the using of bootstrapping methods for regression problems. Since, many real world applications are beginning to appear along with numerous Monte Carlo studies on the performance of the bootstrap and its competitors for various problems. It is also becoming clearer that the bootstrap has significant in practical value but it also has some limitations. A summary of the literature survey on bootstrapping methods and parameter estimation in the area of chemical kinetics is given below.

2.1 REVIEW OF THE PREVIOUS WORK : POINT ESTIMATIONS

Estimation in algebraic equations, which are non-linear in parameters and in differential equations, was reviewed with particular emphasis on application to kinetics, by Seinfeld [39], Bard and Lapidus [1] and Froment and Bischoff [24]. An extensive treatment of the estimation problem in nonlinear regression can be found in the books by Draper and Smith [16a] and Bard Y [2]. Rohatgi [37] and Lehmann [32] have also given useful contributions in this field. Also parameter estimation in design of experiments for model discrimination are available in literature Mezaki et. al. [33], Prasad and Rao [36] etc.

Invariably, in several practical situations, one encounters both single and multi response models for parameter estimation in nonlinear regression. Seber and Wild [38] gave numerical methods to estimate the parameters by least square estimation, one of the well-recognized method is Gauss - Newton method.

Carr[11] studied the catalytic isomerization of n-pentane to 2-methylbutane and the experimental results are analyzed by applying linear least squares regression by transforming the nonlinear models to linear. Jhonson et. al. [29] studied the reexamination of hougen - watson rate models including a weighted least squares approach and found the dual site model is capable for explaining the data than singl site model in the study made by Carr[11].

Another real life data for single response models is studied by Stein [46]. He considered the response model of the polynomial type and parameters estimated at different temperatures.

Buzzi and Forzatti [9] studied model discrimination by parameter estimation in five kinetic models for the synthesis of methanol from carbon monoxide and hydrogen. Their discrimination is based on the values of posterior probabilities for each model.

Hill [26], Singh [45], studied the parameter estimation and model discrimination among exponential models which represents the rates of a chemical reaction. Kanodia [30] recently compared the efficiency of algorithms gauss- newton and simplex methods in parameter estimation of nonlinear models.

Box and Drapper [6] have proposed that for multi response systems, in absence of variance-covariance matrix, minimization of determinant is the ideal objective function. They have added that the least square minimization should be used when the, variance-covariance matrix is known.

Boag [5] *et al.* has used the minimization of the determinant as the objective function for a multi response nonlinear system for the oxidation of a-xylene by vanadia catalyst. The minimization of the determinant was achieved by using Powell's method.

Vajda and Valko [48] have proposed that for multiresponse problems minimization of the determinant, is the objective function. But, they have observed that least square minimization can be used to get good initial guess or starting point

for obtaining the final parameters, which involves minimization of determinant. They have shown that least square minimization can be used in absence of correlation between responses; otherwise the minimization of determinant provides the best estimate of parameters from multi response data.

Buzzi et. al. [10] are proposed the weighted nonlinear regression technique for the estimation of parameters in multi response models which is used in the sequential experimental design for model discrimination.

Recently Negi [35] studied the effect of simplex method for the estimation of parameters in single and multi response models using nonlinear regression.

Moros et. al. [34] described the application of genetic algorithms(GA) for generating the initial parameter estimations for kinetic models of catalytic processes. They also made a study on the rate of convergence of genetic algorithms in parameter estimation.

Bates and Watts [3] fitted the Hougen-Watson model for reaction kinetics by using Gauss – Newton algorithm with Levenberg – Marquardt modifications for global convergence. They showed this combined procedure ensures the global optimum like the other methods which didn't use differentiation for e.g. Simplex method, Powell's method etc.

2.2 REVIEW OF THE PREVIOUS WORK : INTERVAL ESTIMTIONS

Although regression analysis is one of the most widely used statistical techniques, application of the bootstrap to regression problems has only appeared fairly recently. Only a few books like Sen and Srivastava [40], Draper and Smith [16b] and Shao and Tu [43] incorporated a brief discussion of the bootstrap in regression.

Recently Efron and Tibshirini [21] and Chernick [12] gives an updated description of bootstrapping methods on wide range of applications.

Early discussion of the two methods of bootstrapping in the nonlinear regression model can be found in Efron[17] . Efron and Tibshirani [20] provide a variety of interesting applications and some insightful discussion of bootstrap applications in regression problems.

Bootstrapping the residuals is an approach that also can be applied to

nonlinear models. Shimabukuro et. al. [44] presented an early example of a practical application of a nonlinear regression problem. The first major study of the bootstrap as applied to the problem of estimating the standard errors of the regression coefficients by constrained least squares with an unknown, but estimated, residual covariance matrix can be found in Freedman and Peters [23].

Bickel and Freedman [4] also deal with issues regarding bootstrapping in regression problems. Their study is very interesting because it shows that the conventional asymptotic formulas, which are correct for very large samples, do not work well in small to moderate sample size problems. They show that these standard errors can be too small by a factor of nearly three. On the other hand, the bootstrap method gives accurate answers.

Theoretical work on the use of bootstrap in regression is given in Freedman [22], Weber [49], Wu [50] and Shao [41,42]. A subset of articles that present the theoretical justification for the bootstrap are Efron [18,19].

These authors pointed out that there is unfortunately no good rule of thumb to apply to determine when the conventional formulas will work or when it may be necessary to resort to the bootstrap. They suggest that the development of such a rule of thumb could be the result of additional research. Even the bootstrap procedure has problems in this context.

Efron and Tibishirani [21] discuss the methods called the parametric bootstrap and non-parametric bootstrap. The parametric bootstrap is closer to the Monte Carlo methods. They also discussed about bootstrap estimate of bias.

Breimn [8] gave various recent regression applications and also the model selection using bootstrapping methods with the help of Cross validation technique.

Hayes et. al. [25] have extended bootstrap methods to the case of several unrelated samples with application to estimating contrasts in particle physics problems.

2.3 REVIEW OF THE PREVIOUS WORK IN MODEL SELECTION

In many situations involving chemical kinetics the underlying physical theory suggests several rival mechanisms that give rise to different models. The problem becomes one of the choosing the best model from among a group of 'm' rival models.

Cox [13,14] used an hypothesis testing approach to choose between two rival response functions. Cox considered the weighted sum of the functions and tested the null hypothesis versus the alternative hypothesis.

Hunter and Reiner [28] suggested a sequential procedure for design of experiments which leads to model discrimination. They suggested the set of operating conditions that maximized the deviation between the predicted values of the response was chosen for the next run.

Recently, methods like AIC, BIC and EDC are proposed for the choice of a model by minimizing a criterion function defined on the set of alternative models. Hocking [27] and Thomson [47] studied the above different methods in the case linear regression problems. Recently, Kundu D and Murali [31] are discussed about the model selection in linear regression by using Monte Carlo simulation.

2.4 SUMMARY

From the above review, it is clear that parameter estimation and model selection is very important in kinetic modeling. The literature review of interval methods shows the importance of applying them to nonlinear models. To understand the effect of these methods on parameter estimation and cross validation methods for model selection, a simulation study is carried in this thesis. The data is simulated using nonlinear models, which generally observed in chemical reactions. The effect of parameter estimation for different variances is also studied in both single and multi response models. The parameters are estimated by using Gauss-Newton method with Levenberg-Marquardt modifications.

CHAPTER 3

METHODOLOGY

This chapter deals with the techniques and methods used in the parameter stimation and model discrimination for both the single response and the ıultiresponse models. There are two ways in which one can get the data - either using :al life systems or by generating the data by simulation. In the present study the data re generated using simulation. The data simulation is done for single response as vell as for multiresponse example. In addition, a real life example has been onsidered later.

3.1 Single Response Models

For a particular system only one dependent variable is observed then the system is said to be single response system. Suppose y is the observed response, $\underset{\sim}{x}$ is the vector of independent variables and $\underset{\sim}{\beta}$ is the vector of parameters in the equation and y is related to $\underset{\sim}{x}$ and $\underset{\sim}{\beta}$ by function f, then a single response model is given by,

$$y = f(\underset{\sim}{x}, \underset{\sim}{\beta}). \tag{3.1}$$

For example, suppose a chemical reaction is occurring in the following way

$$A \longrightarrow B. \tag{3.2}$$

Depending on the order of the reaction, the expected concentration of A is given by,

$$\eta = \{1 + \theta_1 \xi_1 \exp(-\theta_2 / \xi_2)\}^{-1}, \tag{3.3}$$

where η is expected concentration of A.

θ_1 and θ_2 are the parameters involved in this model.

ξ_1 and ξ_2 are the reaction time and the temperature respectively.

Here equation (3.3) represents a single response model for a chemical reaction used by Hill [26]. This example consists of a nonlinear equation of exponential type. There are various single response models in chemical engineering of different types like polynomials, differential equations and other nonlinear equations.

3.1.1 Simulation of the Data

The data for ξ_1 and ξ_2 and the values of parameters are taken from Hill W.J.[26]. Using these values the theoretical concentration of A are calculated as follows.

$$y = \{1+\theta_1\xi_1 \exp(-\theta_2/\xi_2)\}^{-1} + R^1 \tag{3.4}$$

$$R^1 = R * \sigma + \mu \tag{3.5}$$

where y is the generated value of concentration of A, η is the theoretical (true) value of concentration of A, R^1 is the random number generated at a mean (μ) and variance (σ^2). R is the random number generated from a random population, which is distributed as N (0,1). The response data is generated at different variances by all values of ξ_1 and ξ_2 and the parameters are estimated.

3.1.2 Parameter Estimation

The method of least squares estimation is used to estimate the parameters. The objective function is defined as

$$O.F. = \underset{\theta_1,\theta_2}{Min} \sum_{i=1}^{n} (\{1+\theta_1(\xi_1)_i \exp(-\theta_2/(\xi_2)_i)\}^{-1} - (\eta)_i)^2 . \tag{3.6}$$

Using the simulated values of concentrations, this function is minimized with respect to θ_1 and θ_2 with the help of Guass Newton-Levenberg Marquardt method, which is described in the section 3.3. The 95% confidence intervals for the parameters are calculated by simulations using Asymptotic theory and Bootstrapping methods. The average estimates, mean square errors of the parameters are also calculated. Another real life example is discussed by Carr [11], which involves the parameter estimation by least squares minimization.

3.1.3 Model Selection

Suppose an experimenter is investigating a system in which the underlying theory suggests at the outset that there are number of models which might represent the system. This is not new in chemical reaction engineering. For example, if a chemical reaction is observed as given in equation (3.2), depending on whether the reaction is of first, second, third or forth order, we can consider four rival models. In this case identifying the correct model or model selection plays important role since the correct model gives exact order of the reaction, which is useful to experimenters.

For example, the four rival models for equation (3.2) are given as follows.

$$\eta^{(1)} = \exp\{-\theta_1^{(1)}\xi_1 \exp(-\theta_2^{(1)} / \xi_2)\} \tag{3.7}$$

$$\eta^{(2)} = \{1+\theta_1^{(2)}\xi_1 \exp(-\theta_2^{(2)} / \xi_2)\}^{-1} \tag{3.8}$$

$$\eta^{(3)} = \{1+2\theta_1^{(3)}\xi_1 \exp(-\theta_2^{(3)} / \xi_2)\}^{-1/2} \tag{3.9}$$

$$\eta^{(4)} = \{1+3\theta_1^{(4)}\xi_1 \exp(-\theta_2^{(4)} / \xi_2)\}^{-1/3} \tag{3.10}$$

Hence the model selection plays an important role in data analysis. Here model selection is done with the help of Cross Validation procedure, which is explained in the section 3.5. The data and the rival models which are considered for model selection are given by Hill W.J.[26].

3.2 Multi Response Models

There are some situations in which the experimenter observes more than one response in a system. As in chemical reaction engineering, kinetic investigation frequently involves with the models where more than one rate equation is possible for a reaction. In such situations the experimenter has to observe several responses instead of one and this multi response model can be represented by

$$\underset{\sim}{Y} = \underset{\sim}{f}(\underset{\sim}{X}, \underset{\sim}{\beta}). \tag{3.11}$$

where $\underset{\sim}{Y}$ is the observed response, $\underset{\sim}{X}$ is the vector of independent variables and $\underset{\sim}{\beta}$ is the vector of parameters and f is a nonlinear function. Here the equation for i^{th} response $(1 \le i \le r)$ and u^{th} experiment $(1 \le u \le n)$ is represented by,

$$y_u^{(i)} = \eta_u^{(i)}(\underset{\sim}{X}_u, \underset{\sim}{\beta}). \tag{3.12}$$

For example a multi response model can be represented by the following reaction scheme given in thesis by Negi [35].

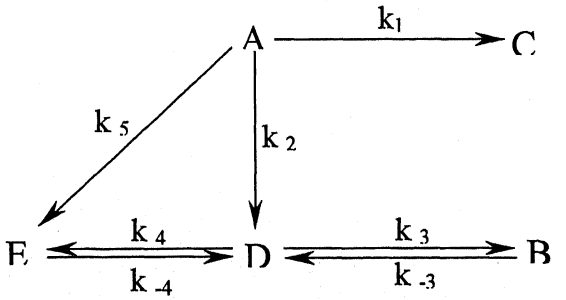


Fig 3.1 Example for a *Multiresponse model*

where A,B,C,D and E denote α-Pinene ($C_{10}H_{15}$), α- and β- Pyronene ($C_{10}H_{16}$), Dependene ($C_{10}H_{16}$), D. *allo*-Ocimene ($C_{10}H_{16}$) and E. Dimer ($C_{20}H_{32}$) respectively and k_1, k_2, k_3, k_{-3}, k_4, k_{-4},k_5 are the reaction rate constants.

The rate equations for the model shown in fig.3.1 are given by,

$$r_A = -(k_1 + k_2)C_A - 2k_5 C_A^2 \tag{3.13}$$

$$r_B = -k_{-3}C_B + k_3 C_B \tag{3.14}$$

$$r_C = k_1 C_A \tag{3.15}$$

$$r_D = k_2 C_A + k_{-3}C_B - k_3 C_D - 2k_4 C_D^2 + 2k_{-4}C_E \tag{3.16}$$

$$r_E = k_5 C_A^2 + k_4 C_D^2 - k_{-4}C_E \tag{3.17}$$

which generally expressed as a multi response system. Here r represents rates of reactions, C represents concentrations of the species and k represents the rate constants, which are unknown parameters.

3.2.1 Simulation of the Data

The theoretical (true) rates are calculated by using the values of independent variables. The response values are generated by adding random numbers to these true values which are generated at different variance-covariances. Let a nonlinear multi response function is given by,

$$\underset{\sim}{Y} = \underset{\sim}{f}(\underset{\sim}{X}, \underset{\sim}{\beta}) + \underset{\sim}{\varepsilon}. \tag{3.18}$$

here $\underset{\sim}{\varepsilon}$ represents the vector of random numbers generated from N(μ, Σ). It is generated as follows.

$$\underset{\sim}{\varepsilon} = Z^T \underset{\sim}{R} + \underset{\sim}{\mu} \tag{3.19}$$

where $\underset{\sim}{R}$ is a vector of standard normal random numbers,

$\underset{\sim}{\mu}$ is a vector representing the mean,

Z^T is a matrix such that $Z^T Z = \Sigma$ and Σ is covariance matrix.

As described in the above procedure multivariate random errors are generated and added to the true values generated by the rate equations.

3.2.2 Parameter Estimation

The parameter estimation in multi response models is not as simple as in the case of single response models. Since the order of the dependent variables is different

for different responses and also there may be correlation between the responses. Hence different procedures are available for parameter estimation in multi response models as seen in literature. The method of weighted least squares estimation is used to estimate the parameters for this type of problems. The parameters are estimated by minimizing the objective function given by,

$$O.F._i = \sum_{u=1}^{N} \left(y_i(X_u) - \hat{y}_i(X_u, \theta_i)\right)^T Q \left(y_i(X_u) - \hat{y}_i(X_u, \theta_i)\right). \qquad (3.20)$$

where Q represents the reciprocal covariance matrix (Σ).

Here also the minimization of objective function is carried with the help of Guass Newton-Levenberg Marquardt method, which is described in the section 3.3. The 95% confidence intervals for the parameters are calculated by simulations using Asymptotic theory and Bootstrapping methods. The average estimates, mean square errors of the parameters are also calculated. A simulated study is made on this estimation as discussed by Buzzi et. al. [10].

3.2.3 Model Selection

It is quite common in chemical engineering practice where more than one chemical reaction takes place and they are typically observed by models that include more than one response. Multi response models which frequently arise in experimental situations contains nonlinear equations and capable of describing the mechanism of the given system equally well. Hence therefore, they are considered to be rival models and the identification of correct model is a challenging problem.

For example, four chemical kinetic models, each having two responses given by Buzzi et. al. [10] are,

$$y_1^{(1)} = \frac{k_{1,1}x_1x_2}{(1+K_{3,1}x_1+K_{4,1}x_2)}, y_2^{(1)} = \frac{k_{2,1}x_1x_2}{(1+K_{3,1}x_1+K_{4,1}x_2)}, \qquad (3.21)$$

$$y_1^{(2)} = \frac{k_{1,2}x_1x_2}{(1+K_{3,2}x_1+K_{4,2}x_2)^2}, y_2^{(2)} = \frac{k_{2,2}x_1x_2}{(1+K_{3,2}x_1)^2}, \qquad (3.22)$$

$$y_1^{(3)} = \frac{k_{1,3}x_1x_2}{(1+K_{3,3}x_1)^2}, y_2^{(3)} = \frac{k_{2,3}x_1x_2}{(1+K_{3,3}x_1)^2}, \qquad (3.23)$$

$$y_1^{(4)} = \frac{k_{1,4}x_1x_2}{(1+K_{3,4}x_1+K_{4,4}x_2)}, y_2^{(4)} = \frac{k_{2,4}x_1x_2}{(1+K_{3,4}x_1)}. \qquad (3.24)$$

Here the model selection is done with the help of Cross Validation procedure, which is explained in the section 3.5. The data and the rival models which are considered for model selection are given by Buzzi et. al. [10].

3.3 Gauss Newton – Levenberg Marquardt Method (GNLM) for Parameter Estimation

This method was first discussed by Bates & Watts [3] and they said it is a powerful one in searching the optimum. This method is nothing but the combination of two methods Guass-Newton & Levenberg-Marquardt. In this method the iterations of the Gauss-Newton method are carried using the Levenberg-Marquardt method to modify the Hessian matrix if the search direction is to far away from the direction of the steepest descent. Hence the failure of the Guass-Newton method will be taken care in most cases where the Hessian matrix is near singular. To understand this method properly first Guass-Newton method and later Levenberg-Marquardt method are explained.

i) GAUSS-NEWTON METHOD

Guass-Newton gave the method as described below.

If the functional form of the model is

$$Y = f(\underline{\xi},\underline{\theta}) + \varepsilon \quad \text{(or)} \quad \eta = E(y) = f(\underline{\xi},\underline{\theta}), \tag{3.25}$$

where $f(\underline{\xi},\underline{\theta})$ is non-linear, then, for a single observation in the k-th experiment, one may write

$$y_k = f(\underline{\xi}_k,\underline{\theta}) + \varepsilon_k, \tag{3.26}$$

and the error sum of squares is given by,

$$S(\underline{\theta}) = \sum_{k=1}^{n} [y_k - f(\underline{\xi}_k, \underline{\theta})]^2 . \tag{3.27}$$

Since y_k and $\underline{\xi}_k$ are known, the sum of squares is a function of $\underline{\theta}$ only. The $\underline{\theta}$ are found such that the sum of squares of errors is a minimum. To find the least squares estimates of $\underline{\theta}$, the function is first expanded about an initial estimate $\hat{\underline{\theta}}_0$ in

Taylor's series and truncated after the first partial derivatives. The linear resultant function is,

$$f(\underline{\xi}_k, \underline{\theta}) = f\left(\underline{\xi}_k, \hat{\underline{\theta}}_0\right) + \sum_{l=1}^{p}\left[\frac{\partial f}{\partial \theta_l}\left(\underline{\xi}_k, \underline{\theta}\right)\right]_{\underline{\theta}=\hat{\underline{\theta}}_0}\left(\theta_l - \hat{\theta}_{l0}\right) \tag{3.28}$$

by abbreviating,

$$f(\theta) \approx f(\theta^{(1)}) + \dot{F}(\theta^{(1)})(\theta - \theta^{(1)}). \tag{3.29}$$

Using this approximation in minimizing the sum of squares with respect to θ, the next approximation of θ is given as

$$\theta^{(a+1)} = \theta^{(1)} + \delta^{(a)}. \tag{3.30}$$

where

$$\delta^{(a)} = \left[\dot{F}(\theta^{(a)})^T F(\dot{\theta}^{(a)})\right]^{-1} \dot{F}(\theta^{(a)})^T (y - f(\theta^{(a)})). \tag{3.31}$$

This provides the iterative scheme of obtaining $\hat{\theta}$. This procedure is iterated until the correction δ becomes exceedingly small.

Seber and Wild [38] also gave a more general approach, viz., the Newton method, which applies to any function satisfying appropriate regularity conditions and in which the residual sum of squares S (θ) is expanded directly using a quadratic Taylor expansion. The minimization of residual sum of squares occurs, when

$$\theta^{(a+1)} = \theta^{(1)} + \delta^{(a)} \tag{3.32}$$

where $\theta^{(1)}$ is the initial guess of estimates and

$$\delta^{(a)} = \left[H(\theta^{(a)})\right]^{-1} J(\theta^{(a)}) \tag{3.33}$$

$$\text{i.e. } H(\theta^{(a)})\delta^{(a)} = J(\theta^{(a)}) \tag{3.34}$$

which can also be rewritten

$$[J(\theta^{(a)})^T J(\theta^{(a)}) + B(\theta^{(a)})]\, \delta^{(a)} = -J(\theta^{(a)}). \tag{3.35}$$

Neglecting the second derivatives matrix $B(\theta^{(a)})$, we obtain the "normal equations" and the Gauss-Newton direction

$$J(\theta^{(a)})^T J(\theta^{(a)})\, \delta^{(k)} = -J(\theta^{(a)}), \tag{3.36}$$

where

$$J(\theta) = \frac{\partial S(\theta)}{\partial \theta} \tag{3.37}$$

and $$H(\theta)=\frac{\partial^2 S(\theta)}{\partial\theta\partial\theta'}. \tag{3.38}$$

The convergence criteria used for halting the procedure is given by

$$\frac{\left|\Delta\hat{\theta}_j\right|}{\left|\hat{\theta}_j\right|}<\gamma, \tag{3.39}$$

where γ is some suitably small value (tolerance) greater than zero.

If convergence is not achieved $\hat{\underline{\theta}}^*$ is updated by replacing the old values by the new values and the process repeated.

ii) LEVENBERG-MARQUARDT METHOD

Levenberg Marquardt gave the method as described below.

Statement of the Problem: The model to be fitted to the data is

$$\begin{aligned} E(y) &= f(x_1,x_2,\ldots,x_k;\theta_1,\theta_2,\ldots,\theta_m) \\ &= f(\underline{x},\underline{\theta}), \end{aligned} \tag{3.40}$$

Where $x_1,x_2,\ldots,x_k$ are independent variables, $\theta_1,\theta_2,\ldots,\theta_m$ are the population values of m parameters, and E (y) is the expected value of the dependent variable y. The data points are denoted by

$$(y_i,x_{1i},x_{2i},\ldots\ldots,x_{ki}), \quad i=1,2,\ldots\ldots,n. \tag{3.41}$$

The problem is to compute those estimates of the parameters that will minimize

$$\phi=\sum_{i=1}^{n}\left[y_i-\hat{y}_i\right]^2=\left\|\underline{y}-\hat{\underline{y}}\right\|^2, \tag{3.42}$$

Where $\hat{y}_i$ is the value of y predicted by (3.41) at the i^{th} data point.

The model is linearized by expanding y_i^* in Taylor series about the current trial values for the coefficients and retaining the linear terms only,

$$y^*_i=\hat{y}^*_i+\left[\frac{\partial\hat{y}_i}{\partial\hat{\theta}_1}\right]^*\Delta\hat{\theta}_1+\left[\frac{\partial\hat{y}_i}{\partial\hat{\theta}_2}\right]^*\Delta\hat{\theta}_2+\ldots\ldots+\left[\frac{\partial\hat{y}_i}{\partial\hat{\theta}_M}\right]^*\Delta\hat{\theta}_M, \tag{3.43}$$

where $\underline{\Delta\hat{\theta}_j} = \left[\hat{\theta}_j - \hat{\theta}_j^*\right]$, j=1,2,.....,M.

The asterisk designates that the quantities are evaluated at the initial trial values. The linearized model is substituted into the objective function and the normal equations formed by setting the partial derivatives of the objective function with respect to each coefficient equal to zero,

$$\frac{\partial \phi}{\partial \hat{\theta}_j} = 0, \qquad j = 1,2,.....,M. \tag{3.44}$$

Thus $\underline{\Delta\hat{\theta}_j}$ can be found by solving

$$A\,\underline{\Delta\hat{\theta}_j} = g \tag{3.45}$$

Where $A^{[m\times x]} = P^T P$,

$$P^{[n\times m]} = \left(\frac{\partial \hat{y}_i}{\partial b_j}\right), \quad i = 1,2...,n;\ j = 1,2,...,m, \tag{3.46}$$

$$g = \left(\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)\frac{\partial \hat{y}_i}{\partial \theta_j}\right). \tag{3.47}$$

The resulting normal equations will be of the form

$$(\underline{J}^T\underline{J} + \lambda_j \underline{I})\underline{\Delta\hat{\theta}_j} = \underline{J}^T(\underline{y} - \hat{\underline{y}}^*), \tag{3.48}$$

Where I is an identity matrix of order m × m and λ_j is a factor that is added to the main diagonal of the $\underline{J}^T\underline{J}$ matrix. The rules for calculating λ_j are discussed in the original article by Marquardt are as follows.

1. Compute $f(\underline{\theta}_j)$.
2. Pick a modest value for λ_j, say $\lambda_j = 0.001$.
3. Solve the linear system of equations (3.48) to find $\underline{\Delta\hat{\theta}_j}$ and calculate $f\left(\underline{\theta}_j + \underline{\Delta\hat{\theta}_j}\right)$.
4. If $f\left(\underline{\theta}_j + \underline{\Delta\hat{\theta}_j}\right) > f(\underline{\theta}_j)$, increase λ_j by a factor of 10, and go back to (3).

5. If $f\left(\underline{\theta}_j + \underline{\Delta\hat{\theta}}_j\right) < f\left(\underline{\theta}_j\right)$, decrease λ_j by a factor of 10, update the trial solution, i.e. replace $\underline{\theta}_j$ by $\underline{\theta}_j + \underline{\Delta\hat{\theta}}_j$ and go to next time step.

The jacobean $\underline{J}$, $\underline{\Delta\theta}$ and error matrices are given by,

$$\underline{J} = \begin{bmatrix} \frac{\partial \hat{y}_1}{\partial \hat{\theta}_1} & \frac{\partial \hat{y}_1}{\partial \hat{\theta}_2} & \cdots & \frac{\partial \hat{y}_1}{\partial \hat{\theta}_M} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial \hat{y}_N}{\partial \hat{\theta}_1} & \frac{\partial \hat{y}_N}{\partial \hat{\theta}_2} & & \frac{\partial \hat{y}_N}{\partial \hat{\theta}_M} \end{bmatrix}^*,$$

$$\underline{\Delta\theta} = \begin{bmatrix} (\hat{\theta}_1 - \hat{\theta}_1^*) \\ (\hat{\theta}_2 - \hat{\theta}_2^*) \\ \cdot \\ \cdot \\ \cdot \\ (\hat{\theta}_M - \hat{\theta}_M^*) \end{bmatrix}, \qquad (\underline{y} - \hat{\underline{y}}^*) = \begin{bmatrix} \left(y_1 - \hat{y}_1^*\right) \\ \left(y_2 - \hat{y}_2^*\right) \\ \cdot \\ \cdot \\ \cdot \\ \left(y_N - \hat{y}_N^*\right) \end{bmatrix}.$$

The $\underline{\Delta\hat{\theta}}$ vector and ϕ will approach zero as convergence is achieved. If convergence is achieved the final coefficients are calculated from

$$\hat{\underline{\theta}}_j = \hat{\underline{\theta}}^* + \underline{\Delta\hat{\theta}}_j, \qquad j = 1,2,\ldots, M.$$

The convergence criteria used for halting the procedure is given by

$$\frac{\left|\underline{\Delta\hat{\theta}}_j\right|}{\left|\hat{\underline{\theta}}_j\right|} < \gamma, \tag{3.49}$$

Where γ is some suitably small value (tolerance) greater than zero.

If convergence is not achieved $\hat{\underline{\theta}}^*$ is updated by replacing the old values by the new values and the process repeated.

In GNLM procedure as said above the modification is just the replacing of equation (3.36) with equation (3.48) which avoids the singularity of $\underline{J}^T\underline{J}$ matrix. A flow sheet illustrating the GNLM procedure is given in Fig. 3.2.

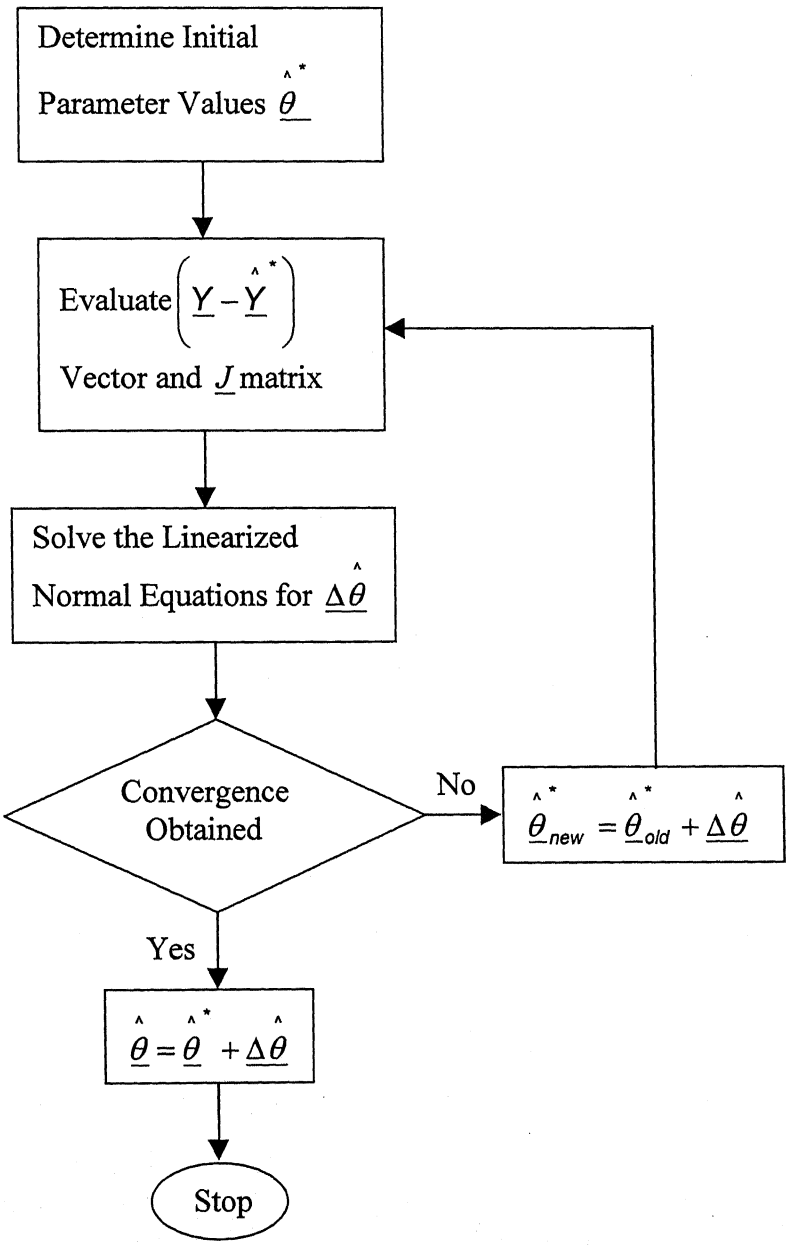


Fig 3.2 *GNLM* Method Logic diagram

3.4 Asymptotic Properties of Non-Linear Least Squares Estimation

Assume the functional form of a nonlinear model is,

$$y_t = f_t(\theta) + \varepsilon_t . \tag{3.50}$$

where t = 1,2,3,…,n.

ε_t are independent identically distributed errors with zero mean and variance $\sigma^2 > 0$.

Then the sum of squared errors function is given by,

$$S(\theta) = \sum_{t=1}^{n} [y_t - f_t(\theta)]^2 . \tag{3.51}$$

The least squares estimate $\hat{\theta}$ is the vector which minimizes equation (3.51) and the solution of $Q^1(\theta) = 0$.

then, $$S'(\hat{\theta}) - S'(\theta) = (\hat{\theta} - \theta) S''(\tilde{\theta}) \tag{3.52}$$

$\tilde{\theta}$ is a point on the line joining $\hat{\theta}$ and θ.

where $S^1(\theta) = \left(\frac{\partial S(\theta)}{\partial \theta_1}, \ldots, \frac{\partial S(\theta)}{\partial \theta_P} \right)$

$$S''(\theta) = \begin{bmatrix} \frac{\partial^2 S(\theta)}{\partial \theta_1^2} \cdots \frac{\partial^2 S(\theta)}{\partial \theta_1 \partial \theta_P} \\ \cdots \\ \cdots \\ \frac{\partial^2 S(\theta)}{\partial \theta_P \partial \theta_1} \cdots \frac{\partial^2 S(\theta)}{\partial \theta_P^2} \end{bmatrix}$$

and $S^1(\hat{\theta}) = 0$.

$$-S'(\theta) = (\hat{\theta} - \theta) S''(\tilde{\theta}) \tag{3.53}$$

$$S''(\theta) = -2\left(\dot{F}(\tilde{\theta}) \dot{F}(\tilde{\theta})^T \right) \tag{3.54}$$

We assume that $[\dot{F}(\tilde{\theta}) \dot{F}(\tilde{\theta})^T]$ is a continuous function of θ and the inverse exists near $\hat{\theta}$ or θ.

$$\therefore S''(\theta) = \frac{\partial^2}{\partial \theta \partial \theta^*} S(\theta) = -2\left[\dot{F}(\theta) \dot{F}(\theta)^T \right]$$

$$S'(\theta) = -2 \sum_{t=1}^{n} \left(y_t - f_t(\theta) \right) \frac{\partial}{\partial \theta} f_t(\theta) = -2\varepsilon \dot{F}(\theta)^T \tag{3.55}$$

The expansion is valid $\forall\ \theta$ in a neighborhood of $\hat{\theta}$ and in particular for θ also for large n.

$$\therefore -\varepsilon\dot{F}(\tilde{\theta})^T = \left(\hat{\theta}-\theta\right)\left[\dot{F}(\tilde{\theta})\dot{F}(\tilde{\theta})^T\right] \tag{3.56}$$

$$(\hat{\theta}-\theta) = -\varepsilon\dot{F}(\tilde{\theta})^T\left[\dot{F}(\tilde{\theta})\dot{F}(\tilde{\theta})^T\right]^{-1} \tag{3.57}$$

$$\Rightarrow \sqrt{n}(\hat{\theta}-\theta) = -\sqrt{n}\left(\varepsilon\dot{F}(\tilde{\theta})^T\left[\dot{F}(\tilde{\theta})\dot{F}(\tilde{\theta})^T\right]^{-1}\right)$$

$$= -\left(\frac{1}{\sqrt{n}}\varepsilon\dot{F}(\tilde{\theta})^T\right)\left(\frac{1}{n}\left[\dot{F}(\tilde{\theta})\dot{F}(\tilde{\theta})^T\right]\right)^{-1} \tag{3.58}$$

as n is large then $\left(\tilde{\theta}\approx\hat{\theta}\approx\theta\right)$ and

$$\left(\frac{1}{n}\left[\dot{F}(\tilde{\theta})\dot{F}(\tilde{\theta})^T\right]\right)^{-1} \to \left(\frac{1}{n}\left[\dot{F}(\theta)\dot{F}(\theta)^T\right]\right)^{-1} \to \Sigma \tag{3.59}$$

where Σ is a positive definite matrix.

$$\sqrt{n}(\hat{\theta}-\theta_0) = -\left(\frac{1}{\sqrt{n}}\varepsilon\dot{F}(\tilde{\theta})^T\right)\left(\frac{1}{n}\left[\dot{F}(\tilde{\theta})\dot{F}(\tilde{\theta})^T\right]\right)^{-1} \tag{3.60}$$

$$\downarrow \qquad\qquad \downarrow$$

$$N_P(0,\sigma^2\Sigma) \qquad \Sigma^{-1}$$

$$\sqrt{n}(\hat{\theta}-\theta_0) = -(X.C) \sim N_P\left(0, C^T D(X) C\right)$$

$$\therefore \sqrt{n}(\hat{\theta}-\theta_0) \simeq N_P\left(0,\sigma^2\Sigma^{-1}\Sigma\Sigma^{-1}\right) = N_P\left(0,\sigma^2\Sigma^{-1}\right) \tag{3.61}$$

where, $$\Sigma = \lim_{n\to\infty}\frac{1}{n}\left(\dot{F}(\theta_0)\dot{F}(\theta_0)^T\right) \tag{3.62}$$

This asymptotic distribution provides us to construct confidence intervals and testing problems.

The necessary assumptions which are required to prove the consistency are,

1. $E(\varepsilon_i) = 0$ and $V(\varepsilon_i) = \sigma^2$, ε_i 's are i.i.d.
2. $f_i(\theta)$ is a continuous function of θ .
3. $\hat{\theta} \in \Theta$, Θ is compact set (closed & bounded). $\hat{\theta}$ is an interior point of Θ.
4. $f(\theta)$ is continuous and bounded for all Θ.
5. Implies that $\frac{1}{n}\sum\left[f(\tilde{\theta}_1)-f(\tilde{\theta}_2)\right]^2$ converges uniformly to $D(\theta_1,\theta_2)$.
6. $\frac{1}{n}\sum\left[f(\tilde{\theta}_1)-f(\tilde{\theta}_2)\right]^2$ converges to $D(\theta_1,\theta_2)$ uniformly and $D(\theta_1,\theta_2) = 0$ iff $\theta_1 = \theta_2$.

3.5 Bootstrapping Methods for Parameter Estimation

The treatment of the bootstrap methods described here comes from Efron and Tibshirani [21]. The interested reader is referred to that text for more information on the underlying theory behind the bootstrap. There does not seem to be a consistent terminology in the literature for what techniques are considered bootstrap methods. Some refer to the resampling techniques of the previous section as bootstrap methods. Here, we use *bootstrap* to refer to Monte Carlo simulations that treat the original sample as the pseudo-population or as an estimate of the population. Thus, in the steps where we randomly sample from the pseudo-population, we now resample from the original sample.

3.5.1 General Bootstrap Methodology

The bootstrap is a method of Monte Carlo simulation where no parametric assumptions are made about the underlying population that generated the random sample. Instead, we use the sample as an estimate of the population. This estimate is called the empirical distribution $\hat{F}$ where each X_i has probability mass $1/n$. Thus, each X_i has the same likelihood of being selected in a new sample taken from $\hat{F}$.

When we use F as our pseudo-population, then we resample *with replacement* from the original sample $x = (x_1,...,x_n)$. We denote the new sample obtained in this manner by $x^* = (x_1^*,...,x_n^*)$. Since we are sampling with replacement from the original sample, there is a possibility that some points X_i will appear more than once in x^* or maybe not at all. We are looking at the univariate situation, but the bootstrap concepts can also be applied in the *d*-dimensional case.

A small example serves to illustrate these ideas. Let's say that our random sample consists of the four numbers x = (5, 8, 3, 2). The following are possible samples x^*, when we sample with replacement from x :

$$X^{*1} = (X_4, X_4, X_2, X_1) = (2,2,8,5)$$
$$X^{*2} = (X_4, X_2, X_3, X_4) = (2,8,3,2).$$

We use the notation X^{*b}, $b = 1,...,$ B for the *b-th* bootstrap data set.

In many situations, the analyst is interested in estimating some parameter e by calculating a statistic from the random sample. We denote this estimate by

$$\hat{\theta} = T = t(x_1, ..., x_n). \tag{3.63}$$

We might also like to determine the standard error in the estimate $\hat{\theta}$ and the bias. The bootstrap method can provide an estimate of this when analytical methods fail. The method is also suitable for situations when the estimator $\hat{\theta} = t(x)$ is complicated.

To get estimates of bias or standard error of a statistic, we obtain B bootstrap samples by sampling with replacement from the original sample. For every bootstrap sample, we calculate the same statistic to obtain the *bootstrap replications* of $\hat{\theta}$, as follows

$$\hat{\theta}^{*b} = t(x^{*b}); \qquad b = 1, ..., B. \tag{3.64}$$

These *B* bootstrap replicates provide us with an estimate of the distribution of $\hat{\theta}$. This is similar to what we did in the previous section, except that we are not making any assumptions about the distribution for the original sample. Once we have the bootstrap replicates in equation (3.64), we can use them to understand the distribution of the estimate.

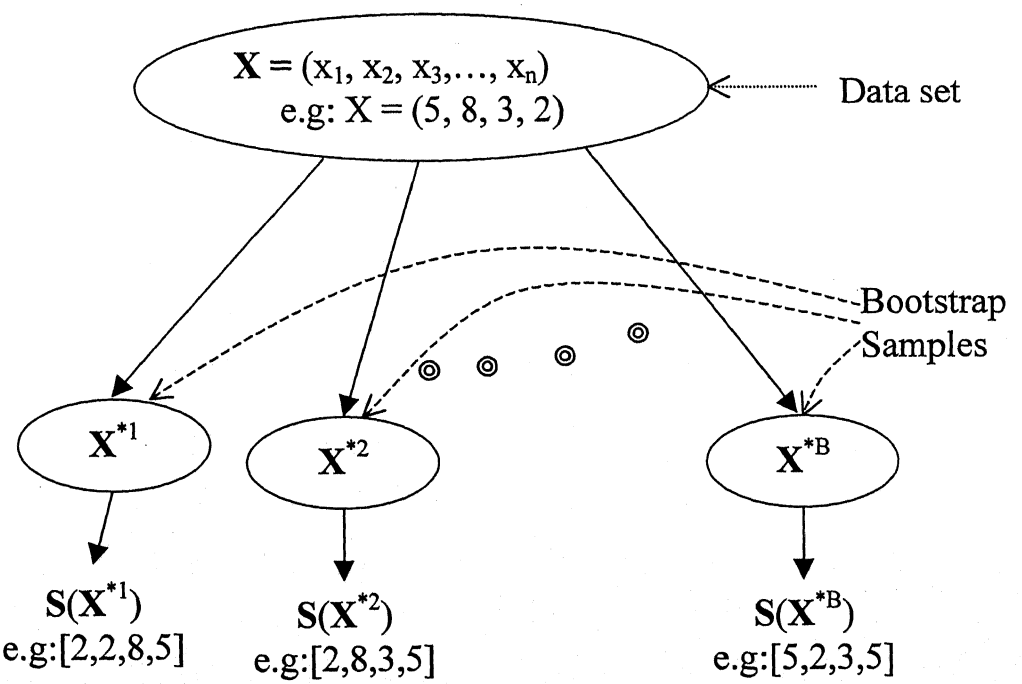


Fig. 3.3 Flow diagram for *Bootstrap Replications*

The steps for the basic bootstrap methodology are given here, with detailed procedures for finding specific characteristics of $\hat{\theta}$ provided later. The issue of how large to make B is addressed with each application of the bootstrap.

PROCEDURE FOR BASIC BOOTSTRAP

1. Given a random sample, $x = (x_1, ..., x_n)$, calculate $\hat{\theta}$.
2. Sample with replacement from the original sample to get $x^* = (x_1^*, ..., x_n^*)$.
3. Calculate the same statistic using the bootstrap sample in step 2 to get, $\hat{\theta}^{*b}$.
4. Repeat steps 2 through 3, B times.
5. Use this estimate of the distribution of $\hat{\theta}$ (i.e., the bootstrap replicates) to obtain the desired characteristic (e.g., standard error, bias or confidence interval).

3.5.2 Bootstrap Percentile Confidence Interval

An improved bootstrap confidence interval is based on the quantiles of the distribution of the bootstrap replicates. This technique has the benefit of being more stable than the *bootstrap-t,* and it also enjoys better theoretical coverage properties (Efron and Tibshirani, [21]). The *bootstrap percentile confidence interval* is

$$(\hat{\theta}_B^{*(\alpha/2)}, \hat{\theta}_B^{*(1-\alpha/2)}), \tag{3.65}$$

where $\hat{\theta}_B^{*(\alpha/2)}$ is the $\alpha/2$ quantile in the bootstrap distribution of $\hat{\theta}^*$. For example, if $\alpha/2 = 0.025$ and B = 1000, then $\hat{\theta}_B^{*(0.025)}$ is the $\hat{\theta}^{*b}$ in the 25th position of the ordered bootstrap replicates. Similarly, $\hat{\theta}_B^{*(0.975)}$ is the replicate in position 975. As discussed previously, some other suitable estimate for the quantile can be used.

The procedure is the same as the general bootstrap method, making it easy to understand and to implement. We outline the steps below.

PROCEDURE FOR BOOTSTRAP PERCENTILE INTERVAL

1. Given a random sample, $x = (x_1, ..., x_n)$, calculate $\hat{\theta}$.
2. Sample with replacement from the original sample to get $x^* = (x_1^*, ..., x_n^*)$.
3. Calculate the same statistic using the bootstrap sample in step 2 to get the bootstrap replicates, $\hat{\theta}^{*b}$.
4. Repeat steps 2 through 3, B times, where $B \geq 1000$.

5. Order the $\hat{\theta}^{*b}$ from smallest to largest.

6. Calculate B. $\alpha/2$ and B. $(1-\alpha/2)$.

7. The lower endpoint of the interval is given by the bootstrap replicate that is in the B. $\alpha/2^{th}$ position of the ordered $\hat{\theta}^{*b}$, and the upper endpoint is given by the bootstrap replicate that is in the B . $(1-\alpha/2)^{th}$ position of the same ordered list. Alternatively, using quantile notation, the lower endpoint is the estimated quantile $\hat{q}_{\alpha/2}$ and the upper endpoint is the estimated quantile $\hat{q}_{1-\alpha/2}$ where the estimates are taken from the bootstrap replicates.

3.6 Cross Validation Method for Model Selection

Often, one of the jobs of a statistician or engineer is to create models using sample data, usually for the purpose of making predictions. We must then decide what model best describes the relationship between the variables and estimate its accuracy.

Unfortunately, in many cases the naive researcher will build a model based on the data set and then use that same data to assess the performance of the model. The problem with this is that the model is being evaluated or tested with data it has already seen. We introduce cross-validation in regression applications, where we are interested in estimating the expected prediction error.

Say we have a set of data, (ξ_i, η_i), where ξ_i denotes a *predictor variable* and η_i represents the corresponding *response variable* of size 'n'. We are interested in modeling the data. Assume there are two competitive models for the available data, which are given by,

$$\eta^{(1)} = \exp\{-\theta_1^{(1)}\xi_1 \exp(-\theta_2^{(1)} / \xi_2)\}, \tag{3.66}$$

$$\eta^{(2)} = \{1+\theta_1^{(2)}\xi_1 \exp(-\theta_2^{(2)} / \xi_2)\}^{-1}. \tag{3.67}$$

The cross validation procedure applied to the above models to select the correct model is explained as follows.

Assume a model and take the given set of data. Remove the i^{th} observation from the data. Then we have (n-1) observations. Estimate the $\hat{\theta}$ from these (n-1)

$$\hat{\eta}_i^{(1)} = \exp\{-\hat{\theta}_1^{(1)}\xi_1 \exp(-\hat{\theta}_2^{(1)} / \xi_2)\} \tag{3.68}$$

Repeat the above procedure for i from 1 to n and add all the squared errors as follows.

$$\sum_{i=1}^{n}(\eta_i - \hat{\eta}_i^{(1)})^2 = CV(1). \tag{3.69}$$

This is Cross Validatory error due to model 1. Repeat the same process for model 2. Then the model which gives minimum value of CV(k) is considered as the best model among the all competing models. If the Cross Validatory error values are nearly equal for different models then further experiments should be done which will discriminate the competing models. This procedure is independent on sample size (n).

PROCEDURE FOR CROSS VALIDATION METHOD

1. Let a Model(k) and sample of size(n).
2. Skip i^{th} Observation from the sample.
3. Estimate parameters using (n-1) Observations.
4. Estimate squared error value for i th Observation using new parameters.
5. Repeat the process for i = 1to n and add all the squared errors as CV(k).
6. The model which gives minimum value of CV(k) is the correct model.

3.7 Estimation of Average Estimates and M.S.E. of the parameters

The average estimates and the mean square errors of the parameters are calculated for the single response and multi response function. The data are simulated in the same fashion as it is in **3.1.1** for single response and in **3.2.1** for multi response example. The work of simulation is repeated 1000 times and the parameters are estimated 1000 times. The average estimates are obtained as follows.

$$Avg.est.of\ k_i = 1/n\sum_{j=1}^{n}(k_i)_j \tag{3.70}$$

The mean square estimate of the parameters is estimated as follows.

$$M.S.E.of\ k_i = 1/n\sum_{j=1}^{n}((k_i)_j - \hat{k}_i)^2 \tag{3.71}$$

CHAPTER 4

RESULTS AND DISCUSSION

The methods described in the last chapter have been applied and the results obtained are given in this chapter.

4.1 Single Response System

The example given in the Section 3.1 is taken and the parameters are estimated for this single response system. The model is

$$\eta = \{1 + \theta_1 \xi_1 \exp(-\theta_2 / \xi_2)\}^{-1}.$$

4.1.1 Simulation of the data

The data are simulated for the single response system. The values of concentrations are taken from Hill [26] and the response values are calculated using the equation

$$\eta = \{1 + 400\xi_1 \exp(-5000 / \xi_2)\}^{-1}. \qquad (4.1)$$

The response values are tabulated as follows.

TABLE 4.1: Single Response data

Run (N)	Input Variables		Response
	(ξ_1)	(ξ_2)	(η)
1	25	575	0.3741
2	25	475	0.7885
3	125	475	0.4271
4	125	575	0.1067
5	125	600	0.0768
6	125	600	0.0768
7	50	450	0.7698
8	100	600	0.0942
9	75	600	0.1217
10	150	550	0.1288
11	25	525	0.5777
12	150	525	0.1856

The errors were generated with different values of the variances and then added to the predicted responses shown in Table 4.1. The generated values of the responses so obtained from Equation (3.4) are shown in Table 4.2.

Table 4.2 The simulated response data with different values of variances

$\sigma = 0$	$\sigma = 0.001$	$\sigma = 0.01$	$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 1$
0.3741	0.3741	0.3525	0.3985	0.2349	-1.4639
0.7885	0.7895	0.7834	0.8558	0.7243	0.8989
0.4272	0.4264	0.4151	0.4285	0.4432	0.0216
0.1068	0.1067	0.1070	0.1490	0.0726	-1.3660
0.0768	0.0779	0.0820	0.2079	0.0573	0.7915
0.0768	0.0787	0.0748	0.0573	0.0372	0.0673
0.7699	0.7717	0.7830	0.7668	0.8507	1.4772
0.0942	0.0937	0.1137	0.1088	-0.0045	-0.8076
0.1218	0.1222	0.1318	0.0835	0.1458	0.1197
0.1288	0.1292	0.1492	0.1334	0.0660	0.2098
0.5777	0.5774	0.5659	0.5190	0.6512	-0.4082
0.1857	0.1857	0.1822	0.2363	0.2655	-0.2455

4.1.2 Parameter Estimation

Using the data in Table 4.1 and the response values generated as shown in Table 4.2 the parameters were estimated using the GNLM method subroutine, given in Appendix 1. For different initial guesses and different variances the estimated parameters are tabulated as follows. The three different initial guesses taken were (410,4900); (400,5000); (390,5100). If we observe the results from Table 4.3 for the three different initial guesses the estimated parameters are same. Hence this assures that we reached the global minima. Also it is clear that the estimated parameters were close to the true values for lower variances.

Table 4.3 Estimated values of the parameters

For σ = 0.001			
	ξ_1	ξ_2	**SOS**
With Initial Guess(1)	399.67	5000.26	9.5547E-4
With Initial Guess(2)	399.67	5000.26	9.5547E-4
With Initial Guess(3)	399.67	5000.26	9.5547E-4
For σ = 0.01			
	ξ_1	ξ_2	**SOS**
With Initial Guess(1)	403.06	4997.65	0.0018
With Initial Guess(2)	403.06	4997.65	0.0018
With Initial Guess(3)	403.06	4997.65	0.0018
For σ = 0.05			
	ξ_1	ξ_2	**SOS**
With Initial Guess(1)	384.24	5010.62	0.0309
With Initial Guess(2)	384.24	5010.62	0.0309
With Initial Guess(3)	384.24	5010.62	0.0309
For σ = 0.1			
	ξ_1	ξ_2	**SOS**
With Initial Guess(1)	415.54	5013.79	0.0587
With Initial Guess(2)	415.54	5013.79	0.0587
With Initial Guess(3)	415.54	5013.79	0.0587
For σ = 1.0			
	ξ_1	ξ_2	**SOS**
With Initial Guess(1)	487.52	4699.04	7.8705
With Initial Guess(2)	487.52	4699.04	7.8705
With Initial Guess(3)	487.52	4699.04	7.8705

True Values: $\xi_1 = 400, \xi_2 = 5000.$

4.1.3 Estimation of Average Estimates and MSE of the Parameters

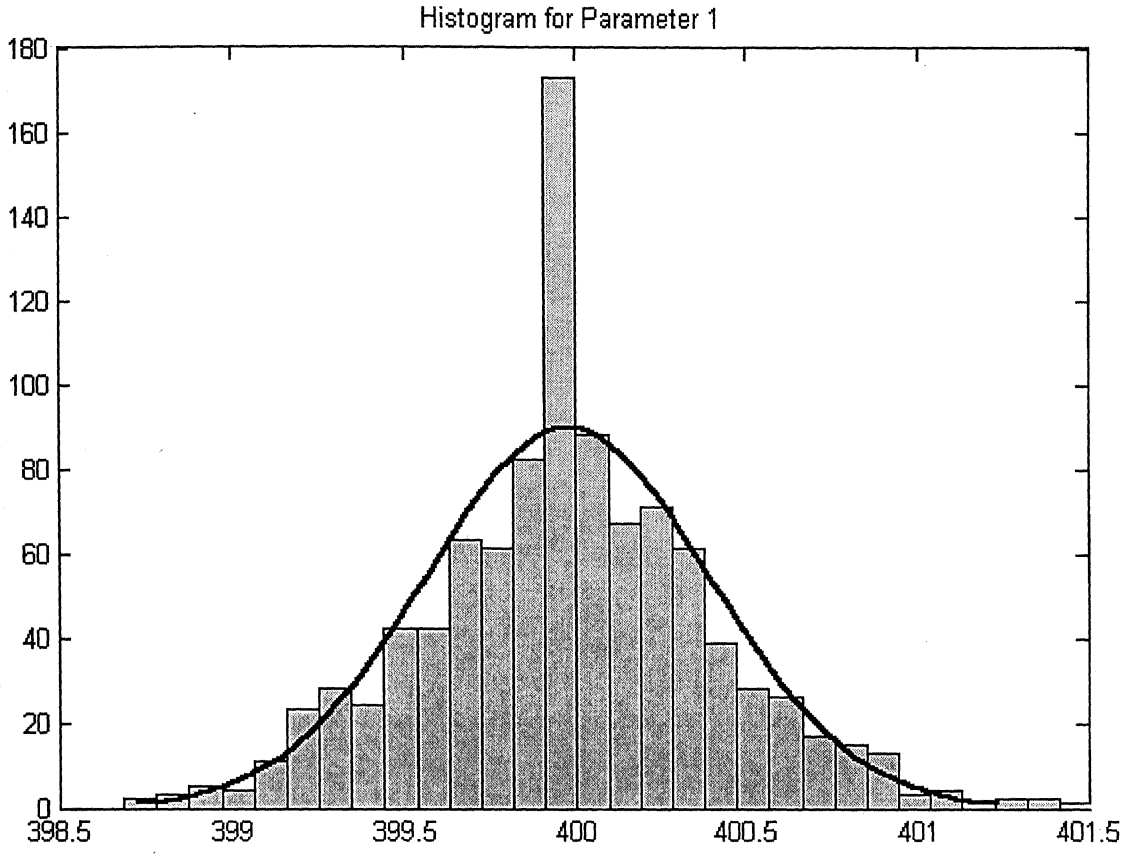
The objective function given by Equation (3.6) is minimized by the above method and repeated for 1000 times. At each time the errors were generated with the same variance and average parameter estimations are calculated. The average estimates and the mean square errors of the parameters were calculated using the Equation (3.70) and (3.71) given in the previous chapter. The values of the average estimates and the mean square errors of the parameters were estimated by taking the different values of variances. The estimated parameters for different values of the variances are given in Table 4.4. However, only for lower values of variances, the estimated parameters were close to the true values. The mean lengths of confidence intervals for the parameters were calculated by using asymptotic theory

and bootstrapping methods given in Table 4.4. The coverage probabilities are also given.

Table 4.4 Simulation Results

σ		**0.001**	**0.01**	**0.05**	**0.1**	**1**
Average Estimates	ξ_1	399.98	399.91	399.54	401.67	429.21
	ξ_2	5000.02	5000.03	4999.85	4988.26	4912.66
Mean Squared Errors	ξ_1	0.1747	17.4880	394.99	1.0561E4	8.4095E4
	ξ_2	0.1069	10.5472	226.51	0.6831E4	9.3212E4
Mean length of asymptotic confidence intervals	ξ_1	32.9570	332.67	1602.92	8615.91	5.9432E4
	ξ_2	42.4182	428.29	2074.82	1.0898E4	9.4086E4
Mean length of bootstrap confidence intervals	ξ_1	1.5229	15.2737	72.5226	346.56	776.56
	ξ_2	1.1711	11.7581	55.0437	279.41	884.78
Coverage probabilities of asymptotic confidence intervals	ξ_1	994	996	995	997	996
	ξ_2	992	993	996	994	992
Coverage probabilities of bootstrap confidence intervals	ξ_1	902	905	899	896	897
	ξ_2	904	905	895	897	891

The average estimates of the parameters gives the more robust values of the estimates, i.e. the values of average estimates do not get change easily by including some more data points. The mean square errors of the parameters give the deviation in the values of estimates with the true values of the parameters. The values of average estimates are again very close to the true values, when the error included in the response decreases. Similarly the values of mean square errors of the parameters are very low when small values of variances used for generating errors. Also the confidence intervals calculated from both the methods are in well agreement. The histograms and the corresponding normality plots of the estimated parameters are given in figures 4.1 & 4.2. From the graphs we can clearly observe that the parameters are normally distributed and the assumptions of Asymptotic theory are satisfied.



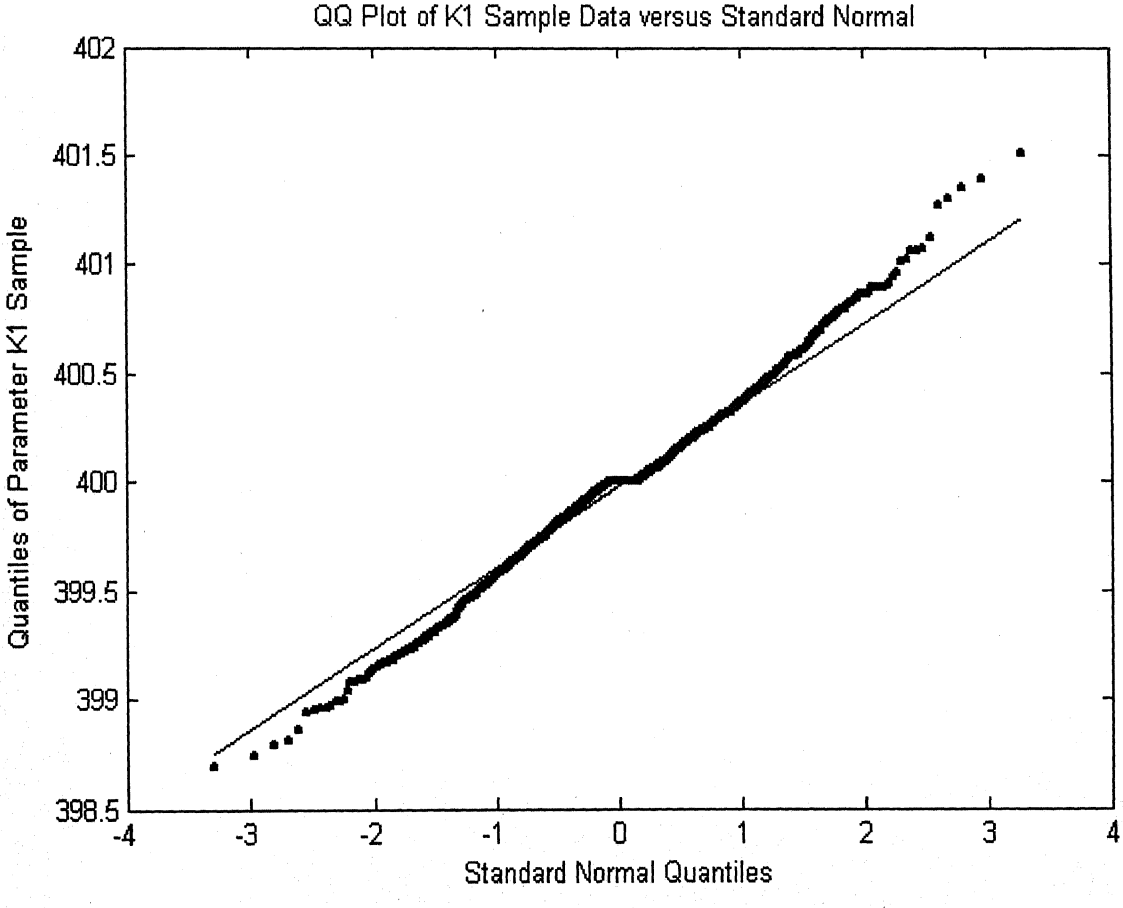
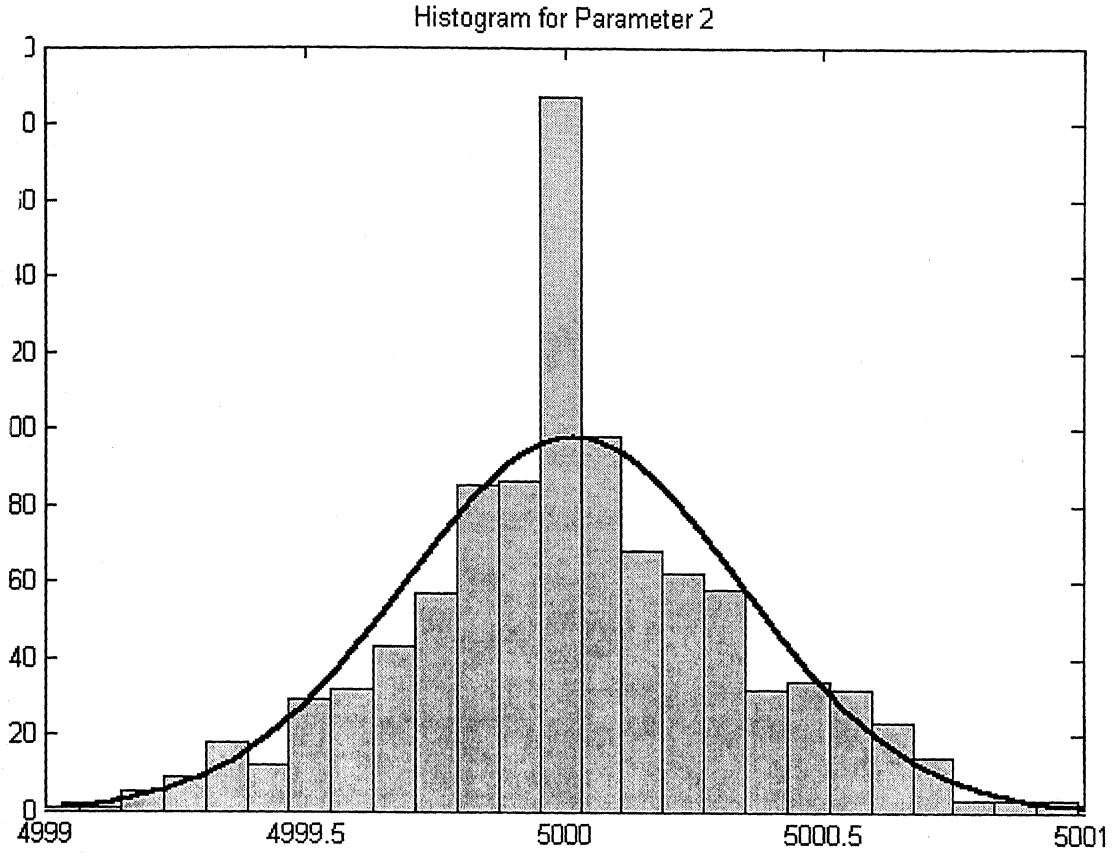


Fig 4.1 Histogram & Q-Q Plot for Parameter ξ_1



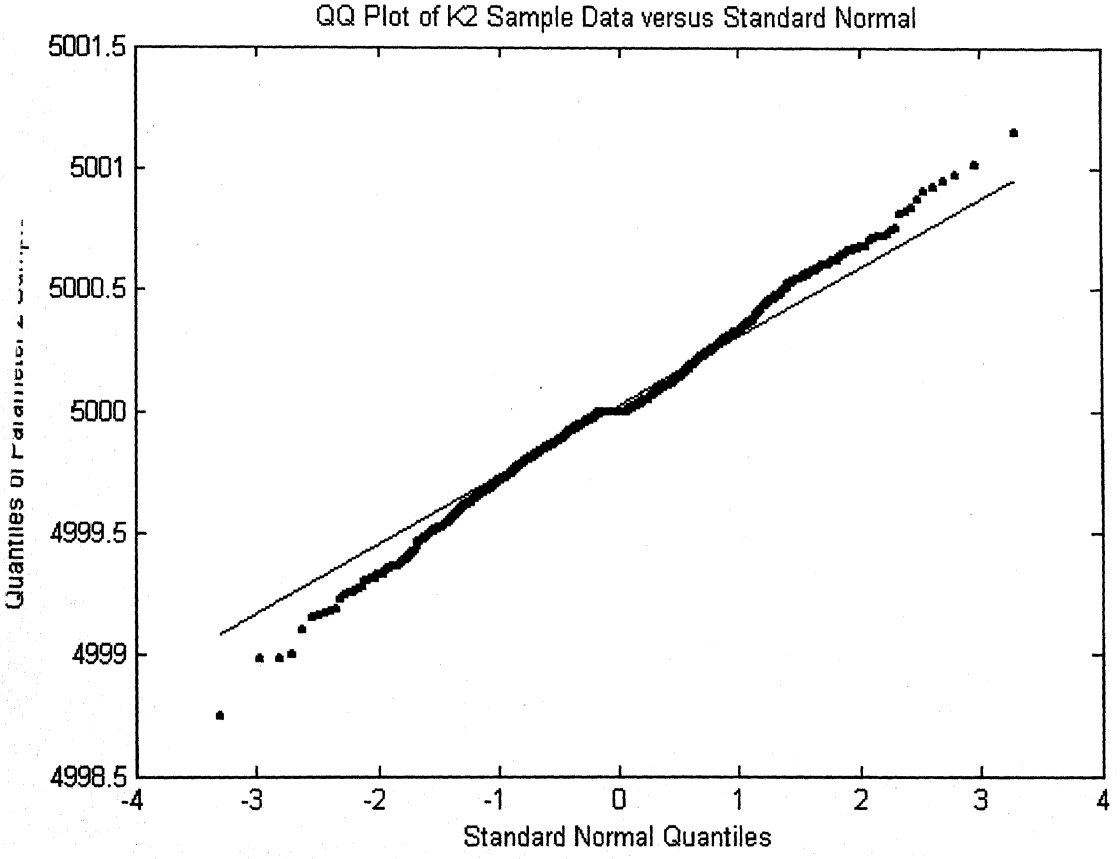


Fig 4.2 Histogram & Q-Q Plot for Parameter ξ_2

.1.4 Model Selection

To see whether the model, from which we generated the data turns out to be rue, the task of model discrimination has been performed. The cross validation nethod is used for model discrimination. The estimates of parameters obtained by ›oth the methods of parameter estimation have been used for model selection. In the :ross validation method the model, having the smallest value of CV is considered as he best one. The value of the cross validation index was calculated using the Equation 3.68 and 3.69. The results are shown in the following Table 4.5.

Table 4.5 Cross Validation Results

σ	CV(1)	CV(2)	CV(3)	CV(4)
0.001	0.0907	1.1485E-5	0.0594	0.1686
0.01	0.1008	0.0022	0.0542	0.1589
0.05	0.1713	0.0355	0.0687	0.1631
0.1	0.1277	0.0728	0.1704	0.3049
1.0	7.5511	6.3935	8.9297	9.2952

From the above results it is evident that for different values of σ, the cross validation index for model 2 is the least among all the rival models. This shows that model 2 is the best among the rival models. Also the studies made by Kanodia [30] says that model 2 is best based on classical methods like *Bayesian* and the *Distance* methods.

4.2 Multi Response System

Buzzi et. al. [10] considered the following chemical kinetic model which is a dual response system.

$$y_1 = \frac{K_1 x_1 x_2}{(1 + K_3 x_1 + K_4 x_2)}, \quad y_2 = \frac{K_2 x_1 x_2}{(1 + K_3 x_1 + K_4 x_2)} \tag{4.2}$$

4.2.1 Simulation of the data

The data were simulated in the same fashion as it was done in single response system. The data were generated from the above model with parameters $K_1 = 0.1$, $K_2 = 0.01$, $K_3 = 0.1$, $K_4 = 0.01$ within operability region $5.0 \le x_1 \le 55.0$, $5.0 \le x_2 \le 55.0$. The response values are tabulated as follows.

TABLE 4.6: Multi Response data

Run (N)	Input Variables		Response	
	(x_1)	(x_2)	(y_1)	(y_2)
1	20.0	20.0	12.5000	1.2500
2	30.0	20.0	14.2857	1.4286
3	20.0	30.0	18.1818	1.8182
4	30.0	30.0	20.9302	2.0930
5	25.0	25.0	16.6667	1.6667
6	25.0	15.0	10.2740	1.0274
7	25.0	35.0	22.7273	2.2727
8	15.0	25.0	13.6364	1.3636
9	35.0	25.0	18.4211	1.8421
10	55.0	32.8	26.4206	2.6421
11	55.0	55.0	42.9078	4.2908
12	10.6	55.0	22.3372	2.2337
13	16.0	55.0	27.9365	2.7937
14	5.0	5.0	1.6129	0.1613
15	5.0	55.0	13.4146	1.3415
16	55.0	5.0	4.1985	0.4198
17	55.0	55.0	42.9078	4.2908
18	10.6	55.0	22.3372	2.2337
19	16.0	55.0	27.9365	2.7937

The data is generated by adding the errors, which are obtained by using different covariance matrices. The simulated data for multi response system are shown in Appendix I.

4.2.2 Parameter Estimation

Using the values of the concentrations and the simulated rates the objective function given by the Equation (3.20) is minimized with respect to the K_1, K_2, K_3 and K_4. The subroutine for GNLM method given in the Appendix II is used for minimization of the objective function with initial guess (1,0.1,1,0.1). The parameters were estimated by taking different values of the variances in simulating the data. The results are tabulated as follows. For lower variances the parameters are near to true values.

TABLE 4.7: Estimated values of the parameters

True values: $K_1 = 0.1$, $K_2 = 0.01$, $K_3 = 0.1$ and $K_4 = 0.01$.

Parameter K_1

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.1036	0.0861	0.1016	0.1038
0.75	0.1050	0.0838	0.1065	0.1078
1.0	0.0975	0.1171	0.0887	0.1134

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0112	0.009	0.0089	0.0122
0.75	0.0113	0.0085	0.0093	0.0088
1.0	0.0101	0.0128	0.0082	0.0144

Parameter K_3

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.1066	0.0878	0.1028	0.1053
0.75	0.1049	0.0807	0.1086	0.1056
1.0	0.0966	0.1145	0.0878	0.1156

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0094	0.0055	0.0097	0.0106
0.75	0.0118	0.0066	0.0117	0.0134
1.0	0.0091	0.0160	0.0083	0.0125

4.2.3 Estimation of Average Estimates and MSE of the Parameters

The objective function given by Equation (3.20) is minimized by the above method and repeated for 1000 times. Each time the error will be generated with the same variance and average parameter estimations are calculated. The average estimates and the mean square errors of the parameters were calculated using the Equation (3.70) and (3.71) given in the previous chapter. The values of the average estimates and the mean square errors of the parameters were estimated by taking the different values of variances. The mean values of estimated parameters for different values of the variances are given in Table 4.8 and their mean squared errors are given in Table 4.9. However, only for lower values of variances, the estimated parameters were close to the true values. The mean lengths of confidence intervals for the parameters are calculated by using asymptotic theory and bootstrapping methods given in Tables 4.10 and 4.11. The coverage probabilities from asymptotic theory and bootstrapping methods are also given in Tables 4.12 and 4.13. The histograms and the corresponding normality plots of the estimated parameters are given figures from 4.3 to 4.14. From the graphs we can clearly observe that the parameters are normally distributed and the assumptions of Asymptotic theory are satisfied. Also the parameters are unbiased for lower variance and higher sample size.

Table 4.9 Estimates of Mean Squared Errors

ılts for Sample Size: 10

Parameter K_1

^{2}	0.5	0.75	1.0	2.0
	2.4621	2.8654	2.5642	
	3.0257	9.4737	9.5911	7.1177
	5.3708	9.995	1.1514	19.3441

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0273	0.0746	0.0827	0.0792
0.75	0.0372	0.1028	0.1061	0.0773
1.0	0.0597	0.1167	0.0153	0.2171

Parameter K_3

$_2^2$	0.5	0.75	1.0	2.0
	2.9116	2.1156	2.5124	3.1956
	3.1823	10.3357	10.4623	8.0099
	5.5289	2.3611	1.1676	21.1755

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.4197	1.3549	1.3958	1.0619
0.75	0.5937	1.5738	1.5941	1.0301
1.0	1.1008	1.4067	0.2631	3.2342

sults for Sample Size: 15

Parameter K_1

$_2^2$	0.5	0.75	1.0	2.0
	1.9311E-4	1.8435E-4	2.095E-4	2.1058E-4
	3.7151E-4	3.2982E-4	3.5005E-4	3.7716E-4
	4.3323E-4	4.3682E-4	4.6771E-4	4.565E-4

Parameter K_2

σ_1^2,σ_2^2	0.5	0.75	1.0	2.0
0.5	0.285E-5	0.35E-5	0.461E-5	0.581E-5
0.75	0.468E-5	0.466E-5	0.52E-5	0.74E-5
1.0	0.523E-5	0.544E-5	0.695E-5	0.846E-5

Parameter K_3

$_2^2$	0.5	0.75	1.0	2.0
	2.4436E-4	2.258E-4	2.345E-4	2.8132E-4
5	4.7141E-5	4.0853E-4	4.3512E-4	4.7867E-4
)	5.4671E-4	5.4527E-4	5.8396E-4	5.7688E-4

Parameter K_4

σ_1^2,σ_2^2	0.5	0.75	1.0	2.0
0.5	0.1427E-4	0.152E-4	0.1651E-4	0.1832E-4
0.75	0.2782E-4	0.2568E-4	0.2618E-4	0.2823E-4
1.0	0.3205E-4	0.3216E-4	0.3633E-4	0.3346E-4

esults for Sample Size: 19

Parameter K_1

σ_2^2	0.5	0.75	1.0	2.0
5	1.5612E-4	1.4653E-4	1.685E-4	1.7831E-4
75	2.6082E-4	2.41E-4	2.7996E-4	2.6434E-4
0	3.5555E-4	3.5793E-4	3.4903E-4	3.557E-4

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.206E-5	0.254E-5	0.261E-5	0.28E-5
0.75	0.312E-5	0.31E-5	0.379E-5	0.484E-5
1.0	0.412E-5	0.446E-5	0.457E-5	0.569E-5

Parameter K_3

$,\sigma_2^2$	0.5	0.75	1.0	2.0
.5	1.8509E-4	1.904E-4	2.0584E-4	2.1163E-4
75	3.0875E-4	2.8455E-4	3.3408E-4	3.129E-4
.0	4.2172E-4	4.2341E-4	4.126E-4	4.2138E-4

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.1227E-4	0.1347E-4	0.128E-4	0.1537E-4
0.75	0.202E-4	0.1873E-4	0.2093E-4	0.2047E-4
1.0	0.2702E-4	0.261E-4	0.2695E-4	0.2803E-4

Table 4.10 Estimates of Mean Length of Asymptotic Confidence Intervals

ults for Sample Size: 10

Parameter K_1

σ_2^2	0.5	0.75	1.0	2.0
5	0.2066	0.1647	0.1767	0.2248
75	38.4261	129.4084	123.1161	100.2374
0	74.9554	71.3931	47.1446	383.2195

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0223	0.0178	0.0193	0.0247
0.75	4.2569	13.4259	12.9731	10.4108
1.0	8.2075	8.0480	4.7599	40.1699

Parameter K_3

σ_2^2	0.5	0.75	1.0	2.0
.5	0.218	0.1764	0.1891	0.2398
75	39.4164	134.0697	128.2142	107.0142
.0	86.1908	82.2281	48.0341	397.5716

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0879	0.0704	0.0758	0.0974
0.75	17.0155	52.9719	48.6026	37.1306
1.0	23.5093	21.8926	20.2096	157.0165

sults for Sample Size: 15

Parameter K_1

$, \sigma_2^2$	0.5	0.75	1.0	2.0
.5	0.0553	0.0615	0.0665	0.0888
.75	0.0645	0.0695	0.0755	0.0981
.0	0.0697	0.0739	0.0819	0.1021

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0066	0.0073	0.0079	0.0107
0.75	0.0076	0.0082	0.009	0.0116
1.0	0.0082	0.0087	0.0097	0.0122

Parameter K_3

$, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.0619	0.0688	0.0745	0.0994
.75	0.0722	0.0777	0.0845	0.1098
1.0	0.078	0.0828	0.0917	0.1143

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0153	0.0171	0.0185	0.0247
0.75	0.0179	0.0192	0.0209	0.0272
1.0	0.0193	0.0205	0.0227	0.0283

esults for Sample Size: 19

Parameter K_1

$^2, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.0488	0.055	0.0589	0.077
0.75	0.056	0.0607	0.0669	0.0845
1.0	0.0618	0.0650	0.0702	0.0890

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0057	0.0064	0.0069	0.009
0.75	0.0065	0.007	0.0077	0.0098
1.0	0.0072	0.0075	0.0082	0.0103

Parameter K_3

$^2, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.0532	0.0599	0.0642	0.0839
0.75	0.061	0.0662	0.0729	0.092
1.0	0.0675	0.0709	0.0765	0.0970

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0135	0.0152	0.0163	0.0213
0.75	0.0155	0.0168	0.0185	0.0233
1.0	0.0170	0.0180	0.0194	0.0246

Table 4.11 Estimates of Mean Length of Bootstrap Confidence Intervals

ılts for Sample Size: 10

Parameter K_1

σ_2^2	0.5	0.75	1.0	2.0
	4.9769	4.6643	4.5620	3.9466
5	9.1990	11.11	10.6483	12.0243
)	15.3921	16.1539	15.2691	16.2877

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.4933	0.4651	0.4503	0.3855
0.75	0.9027	1.1088	1.0436	1.1666
1.0	1.5293	1.6241	1.5507	1.6510

Parameter K_3

σ_2^2	0.5	0.75	1.0	2.0
5	5.5242	5.1866	5.0869	4.3749
'5	10.1823	12.3033	11.7022	13.1788
0	16.9915	17.7304	16.6584	17.9904

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	1.7845	1.6541	1.5757	1.3743
0.75	3.3336	4.1029	4.0012	4.5318
1.0	5.9110	6.0993	5.8796	6.2680

sults for Sample Size: 15

Parameter K_1

, σ_2^2	0.5	0.75	1.0	2.0
.5	0.1382	0.0851	0.1333	0.0937
.75	0.319	0.4481	0.4238	0.3457
.0	0.6211	0.9057	0.9622	0.7749

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0145	0.0096	0.0153	0.0114
0.75	0.0319	0.0442	0.041	0.0357
1.0	0.0629	0.0936	0.0945	0.0731

Parameter K_3

, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.1640	0.1003	0.1596	0.1109
.75	0.3818	0.544	0.5043	0.4147
1.0	0.7583	1.0879	1.1732	0.9507

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0362	0.0234	0.0356	0.0256
0.75	0.083	0.1137	0.1103	0.0898
1.0	0.1566	0.2273	0.2359	0.1871

tesults for Sample Size: 19

Parameter K_1

$_1^2, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.0580	0.0585	0.0568	0.077
0.75	0.08	0.0756	0.1432	0.0811
1.0	0.1525	0.0957	0.2342	0.2613

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0065	0.0069	0.0070	0.009
0.75	0.0086	0.0084	0.0136	0.0103
1.0	0.0164	0.0103	0.0257	0.0233

Parameter K_3

$_1^2, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.0658	0.0666	0.0647	0.0839
0.75	0.0912	0.0857	0.1657	0.0925
1.0	0.1782	0.1094	0.2641	0.2472

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0155	0.0155	0.0151	0.0213
0.75	0.0211	0.0201	0.036	0.0213
1.0	0.0405	0.0253	0.0641	0.0554

Table 4.8 Mean Parameter Estimates

ıe values: **K1 = 0.1** **K2=0.01** **K3 = 0.1** **K4 = 0.01**

ılts for Sample Size: 10

Parameter K_1

σ_2^2	0.5	0.75	1.0	2.0
5	0.1082	0.1131	0.1132	0.1135
75	0.1599	0.3086	0.311	0.3137
0	0.2649	0.2603	0.4373	0.5141

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0108	0.0114	0.0114	0.0115
0.75	0.0163	0.0313	0.0317	0.0323
1.0	0.0279	0.0279	0.0449	0.0541

Parameter K_3

σ_2^2	0.5	0.75	1.0	2.0
.5	0.1085	0.1137	0.1139	0.1142
75	0.1668	0.3181	0.3206	0.3234
.0	0.2877	0.2822	0.4535	0.533

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0135	0.0153	0.0153	0.0154
0.75	0.0309	0.094	0.0949	0.0961
1.0	0.0635	0.0618	0.1454	0.1757

ɛsults for Sample Size: 15

Parameter K_1

$, \sigma_2^2$	0.5	0.75	1.0	2.0
).5	0.1017	0.1021	0.1021	0.1021
.75	0.1026	0.1042	0.1031	0.1032
1.0	0.1036	0.1022	0.1042	0.1043

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0101	0.0103	0.0103	0.0103
0.75	0.0103	0.0104	0.0104	0.0104
1.0	0.0103	0.0102	0.0105	0.0106

Parameter K_3

$^2, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.1019	0.1023	0.1022	0.1023
).75	0.1028	0.1047	0.1034	0.1035
1.0	0.1039	0.1025	0.1046	0.1047

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0104	0.0106	0.0106	0.0106
0.75	0.0107	0.0112	0.0109	0.0109
1.0	0.011	0.0106	0.0112	0.0112

Results for Sample Size: 19

Parameter K_1

$^2, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.1011	0.1018	0.1018	0.1018
0.75	0.1023	0.1021	0.1026	0.1026
1.0	0.1031	0.1018	0.1028	0.1034

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0101	0.0102	0.0102	0.0102
0.75	0.0103	0.0102	0.0103	0.0103
1.0	0.0104	0.0102	0.0103	0.0103

Parameter K_3

$_1^2, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	0.1012	0.1019	0.1019	0.1019
0.75	0.1026	0.1022	0.1028	0.1028
1.0	0.1035	0.102	0.1029	0.1036

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	0.0103	0.0105	0.0105	0.0105
0.75	0.0106	0.0106	0.0108	0.0108
1.0	0.0108	0.0105	0.0108	0.011

Table 4.12 Coverage Probabilities of Asymptotic Confidence Intervals

ults for Sample Size: 10

Parameter K_1

σ_2^2	0.5	0.75	1.0	2.0
5	919	940	959	971
5	886	910	910	960
)	879	906	901	940

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	925	937	948	952
0.75	899	919	911	948
1.0	898	920	914	930

Parameter K_3

σ_2^2	0.5	0.75	1.0	2.0
5	933	935	964	972
'5	895	922	908	960
0	878	908	900	944

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	914	940	958	966
0.75	894	908	923	958
1.0	881	904	912	938

sults for Sample Size: 15

Parameter K_1

σ_2^2	0.5	0.75	1.0	2.0
.5	951	962	975	995
75	925	939	947	980
.0	917	922	948	970

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	951	951	948	975
0.75	938	935	955	955
1.0	934	937	950	962

Parameter K_3

σ_2^2	0.5	0.75	1.0	2.0
.5	948	960	972	995
75	924	938	951	979
.0	912	921	942	964

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	953	960	973	997
0.75	924	943	950	984
1.0	917	924	948	974

sults for Sample Size: 19

Parameter K_1

$, \sigma_2^2$	0.5	0.75	1.0	2.0
.5	951	960	964	985
.75	929	941	955	977
.0	909	927	936	975

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	947	952	958	952
0.75	938	957	953	961
1.0	926	938	930	955

Parameter K_3

$, \sigma_2^2$	0.5	0.75	1.0	2.0
0.5	950	964	965	985
.75	927	942	956	980
1.0	907	923	940	971

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	951	962	970	990
0.75	921	945	962	977
1.0	914	918	935	978

Table 4.13 Coverage Probabilities of Bootstrap Confidence Intervals

Results for Sample Size: 10

Parameter K_1

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	933	926	943	934
0.75	939	933	932	933
1.0	927	943	934	934

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	933	925	932	942
0.75	931	933	928	943
1.0	927	942	944	924

Parameter K_3

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	941	928	943	934
0.75	946	942	928	939
1.0	925	944	934	941

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	908	904	919	926
0.75	906	910	920	919
1.0	906	921	927	908

Results for Sample Size: 15

Parameter K_1

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	937	950	936	950
0.75	942	940	929	943
1.0	950	937	949	938

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	931	932	940	945
0.75	940	929	941	935
1.0	939	936	951	934

Parameter K_3

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	938	948	937	949
0.75	941	940	935	941
1.0	948	937	949	940

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	955	964	955	959
0.75	956	947	951	957
1.0	962	954	961	956

Results for Sample Size: 19

Parameter K_1

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	936	926	925	936
0.75	924	932	920	925
1.0	926	923	934	935

Parameter K_2

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	929	914	940	909
0.75	924	938	927	927
1.0	919	939	924	928

Parameter K_3

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	933	925	929	941
0.75	922	920	920	921
1.0	925	929	940	933

Parameter K_4

σ_1^2, σ_2^2	0.5	0.75	1.0	2.0
0.5	935	932	930	941
0.75	917	935	932	919
1.0	931	932	942	933

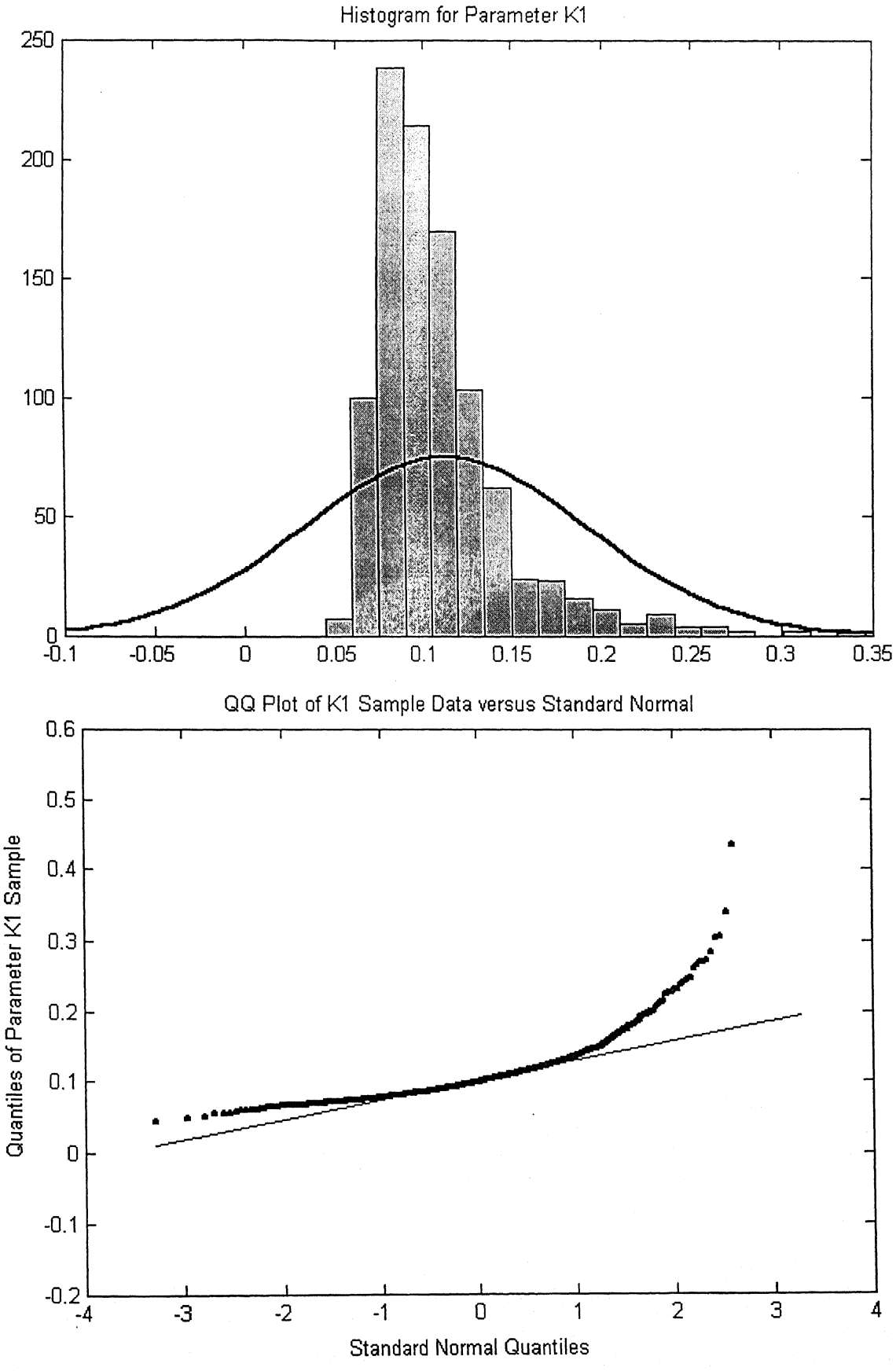
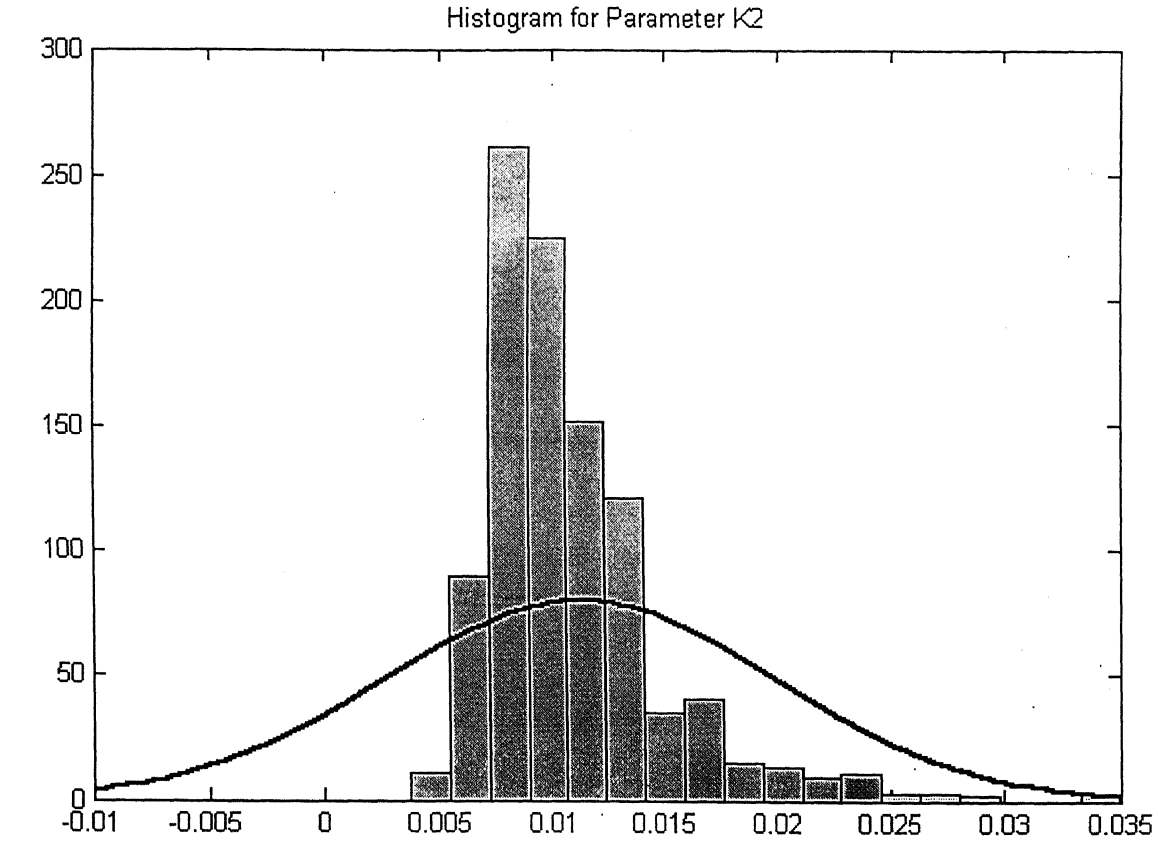


Fig 4.3 Histogram & Q-Q Plot for Parameter K_1 (n=10)



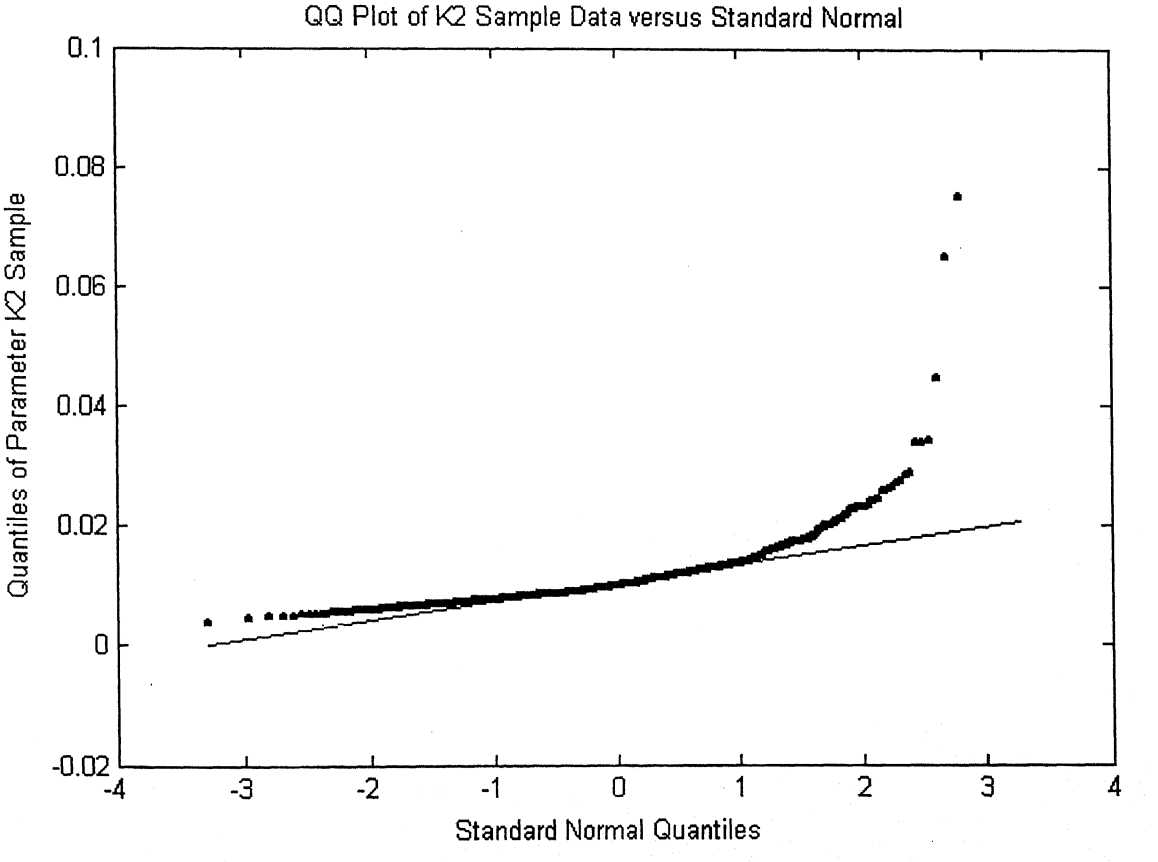
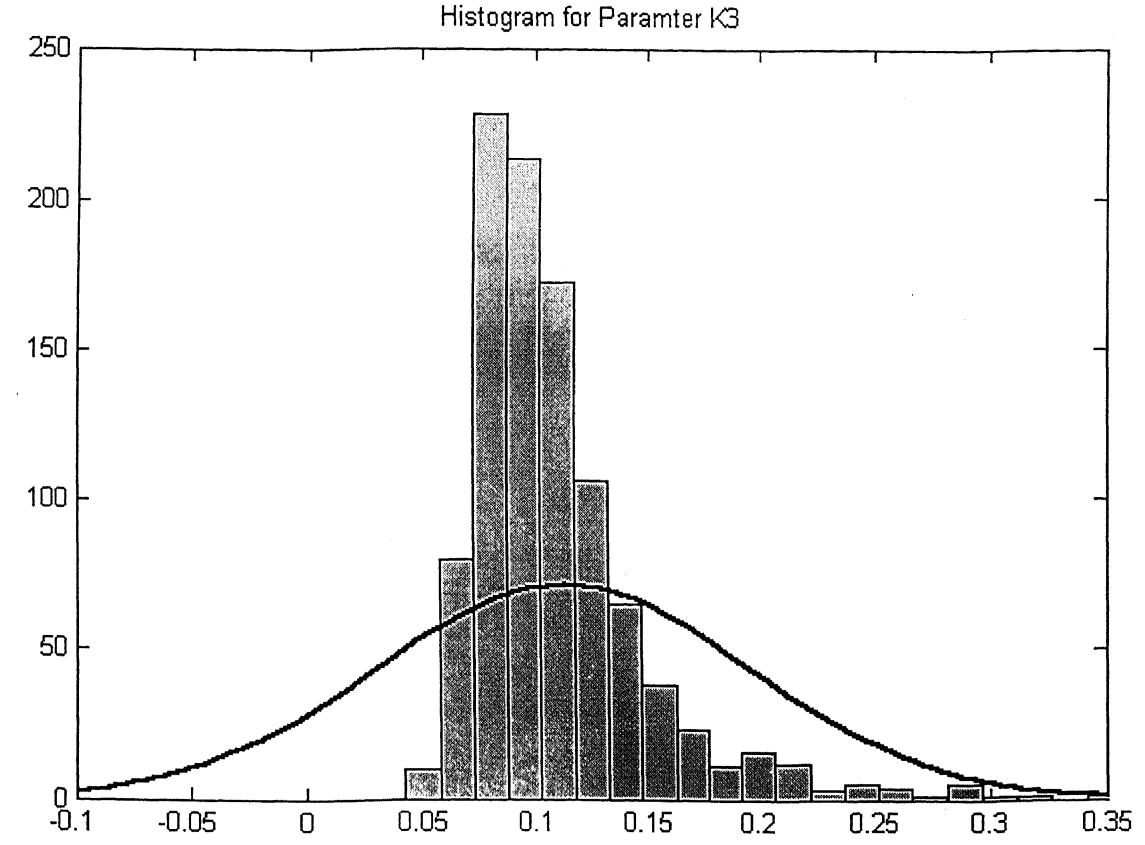


Fig 4.4 Histogram & Q-Q Plot for Parameter K_2 (n=10)



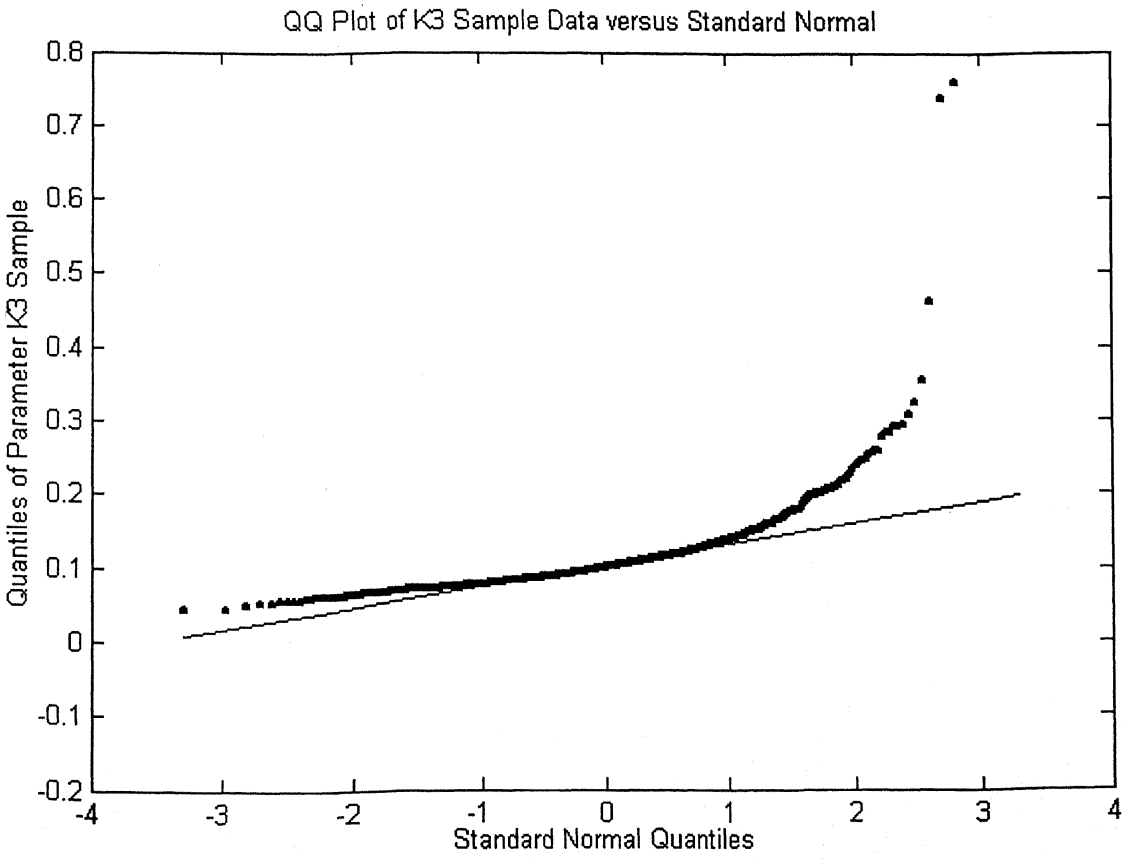


Fig 4.5 Histogram & Q-Q Plot for Parameter K_3 (n=10)

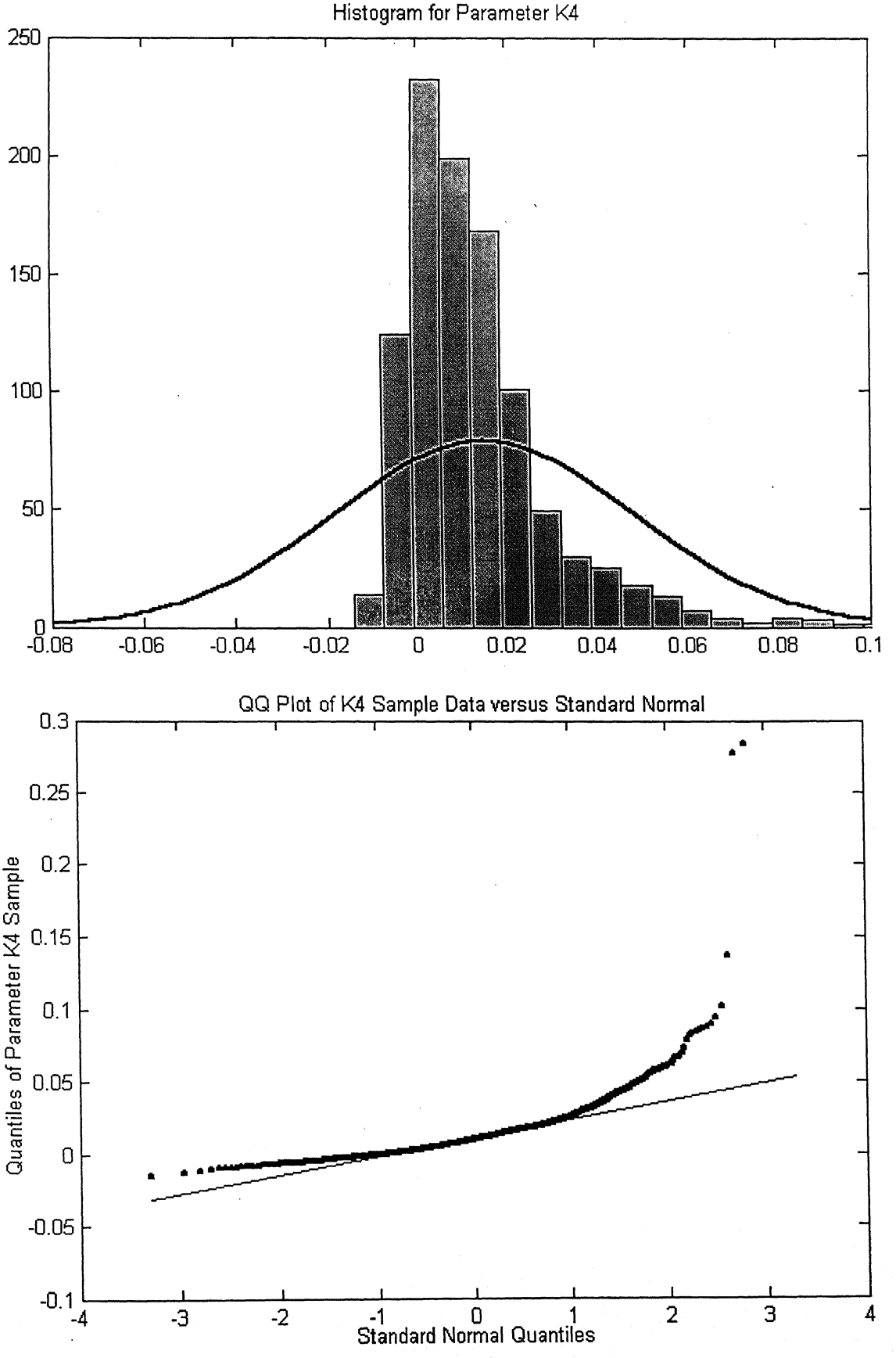
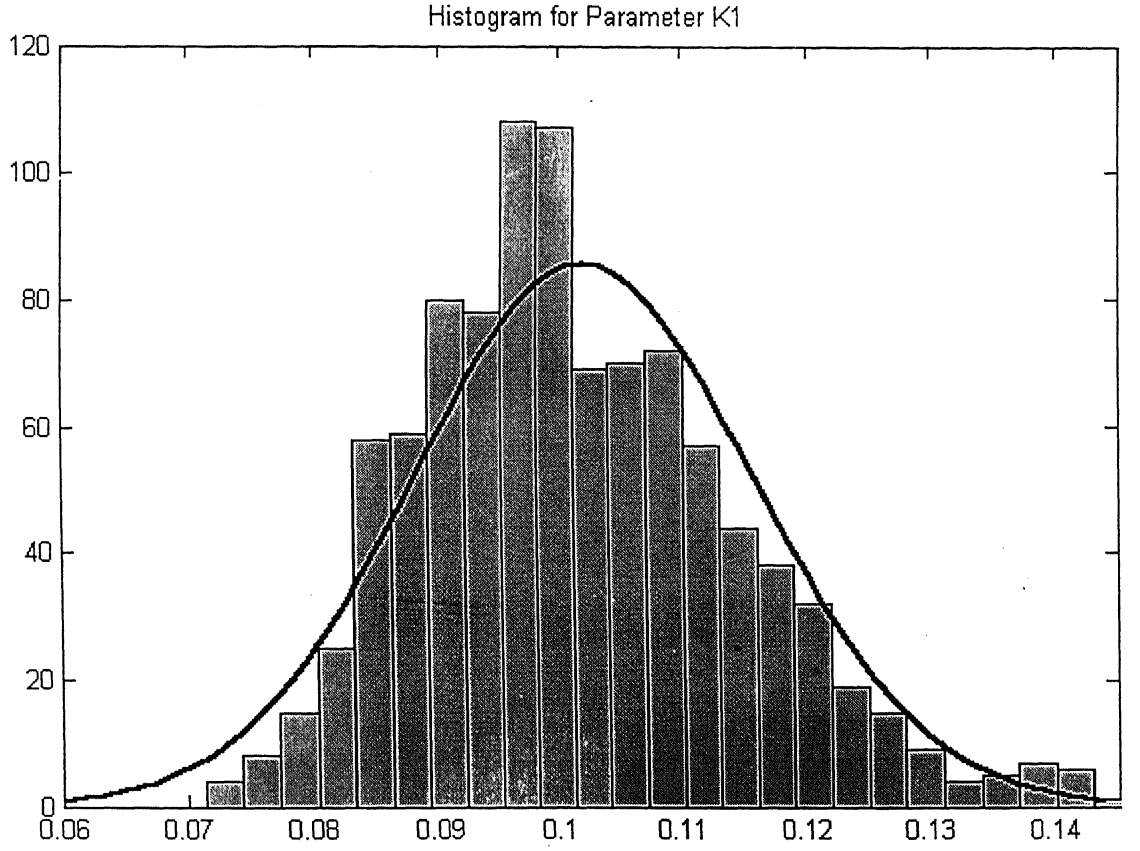


Fig 4.6 Histogram & Q-Q Plot for Parameter K_4 (n=10)



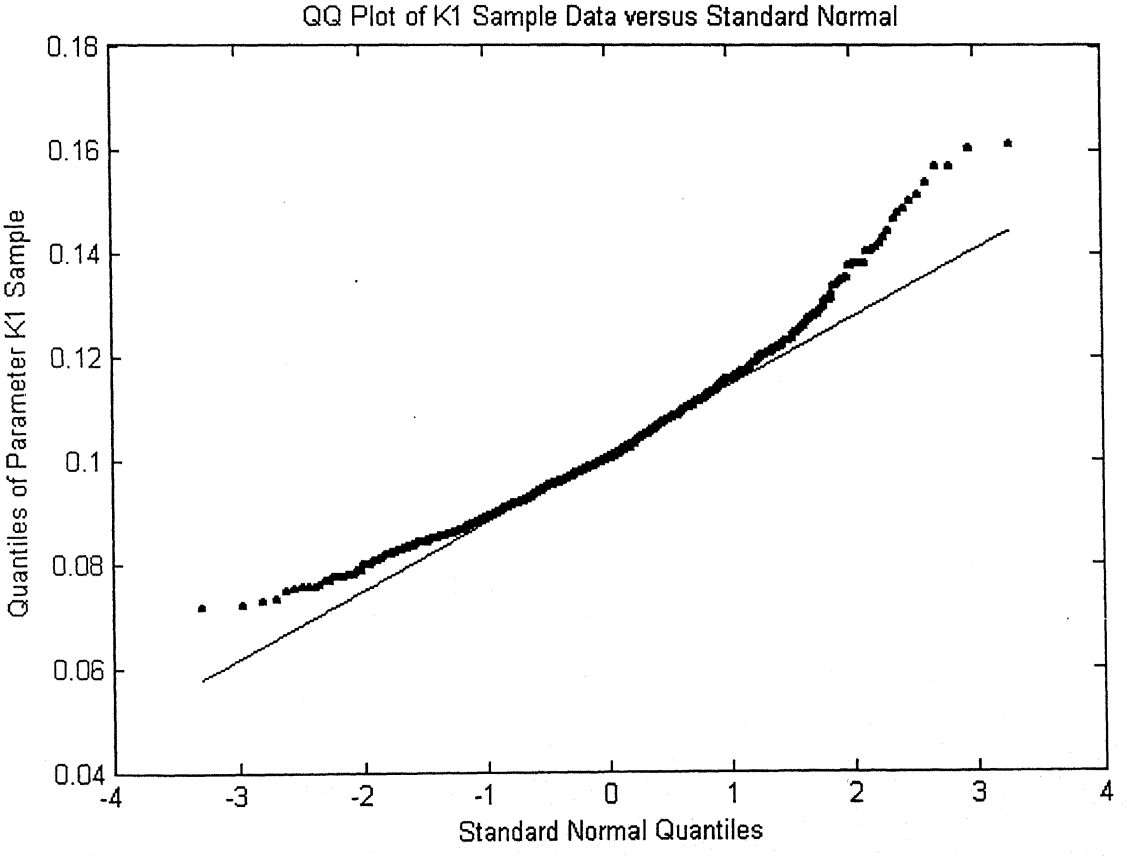
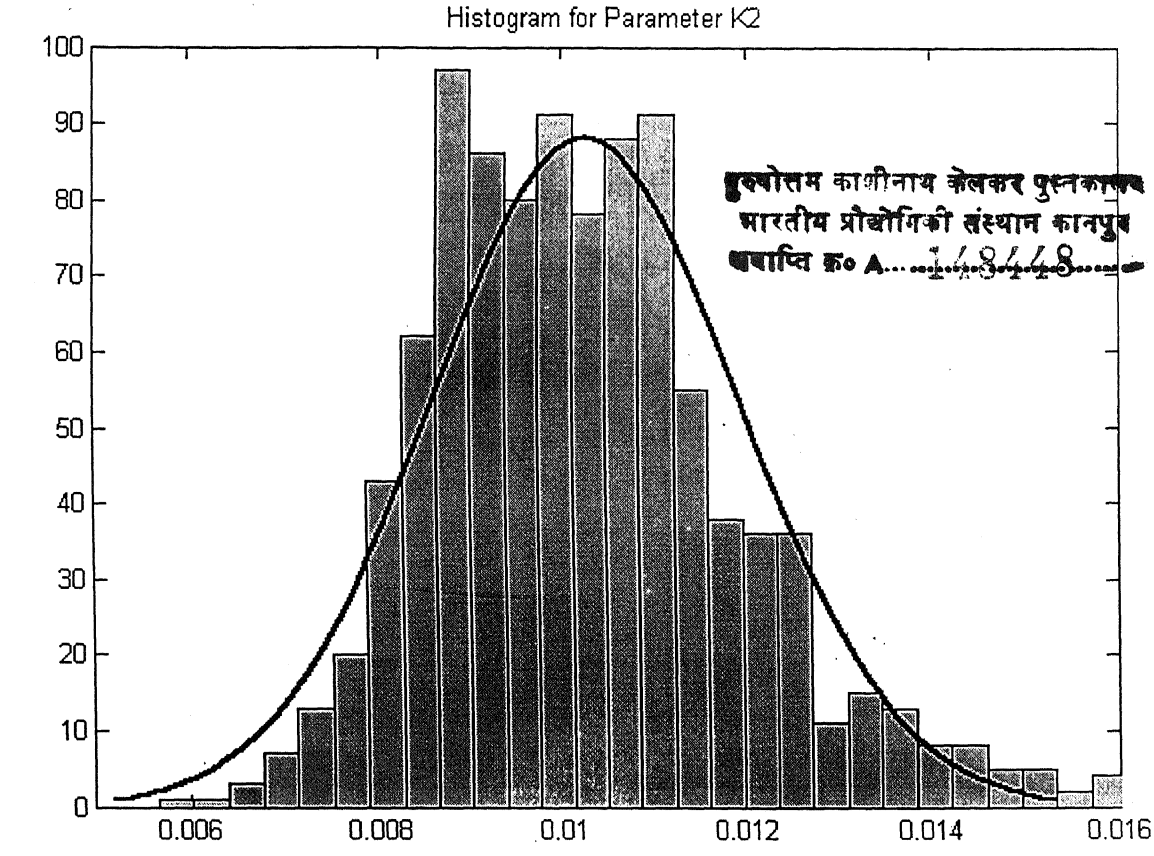


Fig 4.7 Histogram & Q-Q Plot for Parameter K_1 (n=15)



पुरुषोत्तम कार्तीनाथ केलकर पुस्तकालय
भारतीय प्रौद्योगिकी संस्थान कानपुर
प्राप्ति क्र० A...148448...

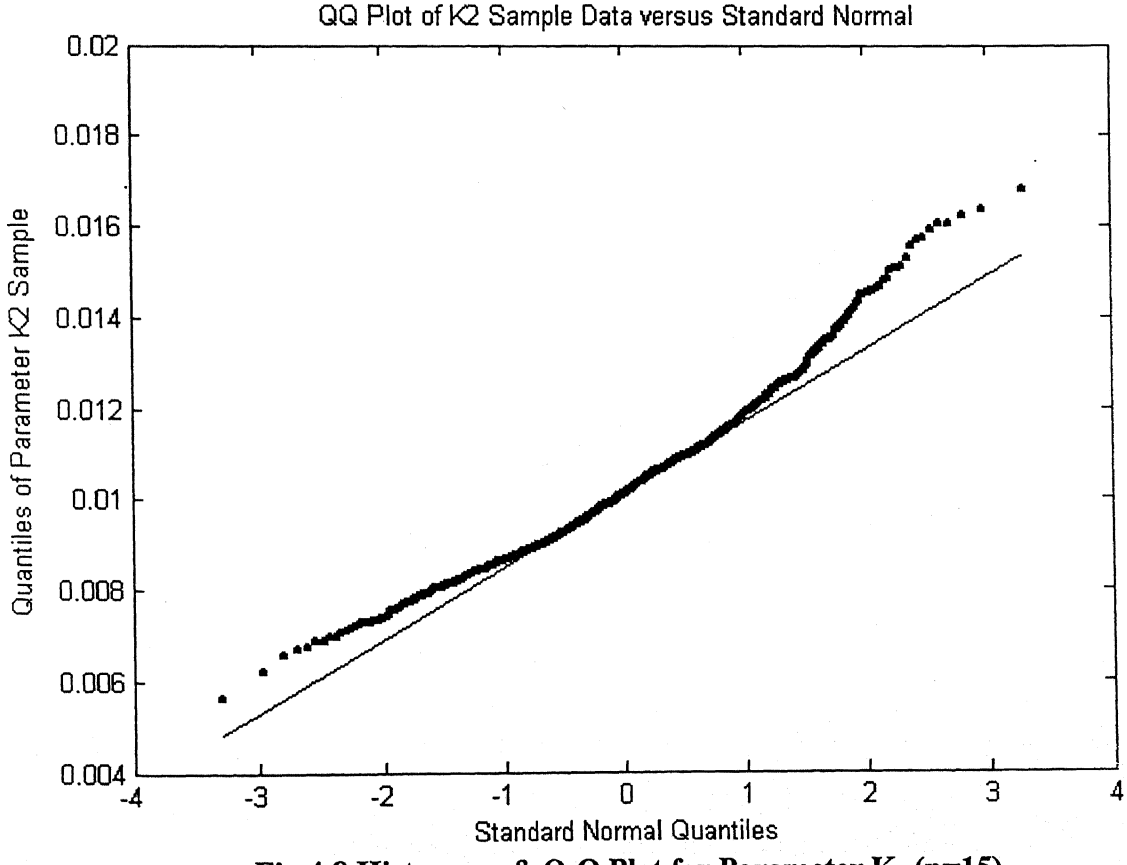
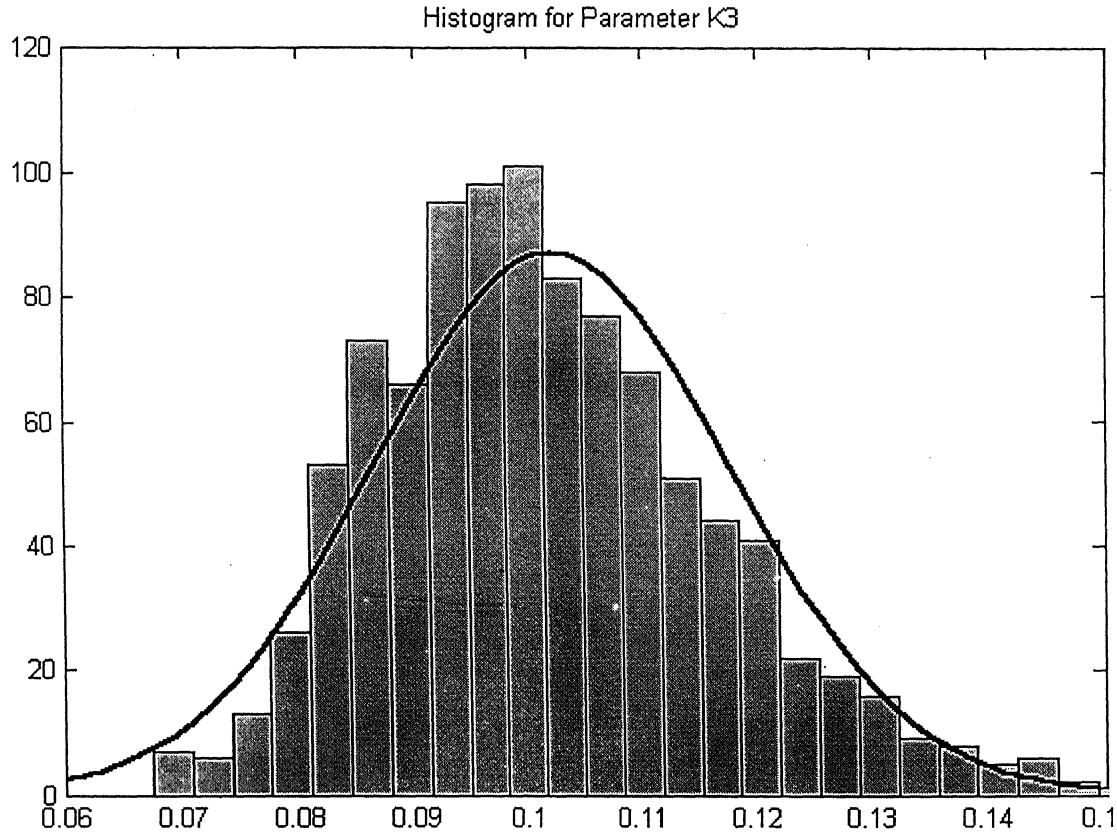


Fig 4.8 Histogram & Q-Q Plot for Parameter K_2 (n=15)



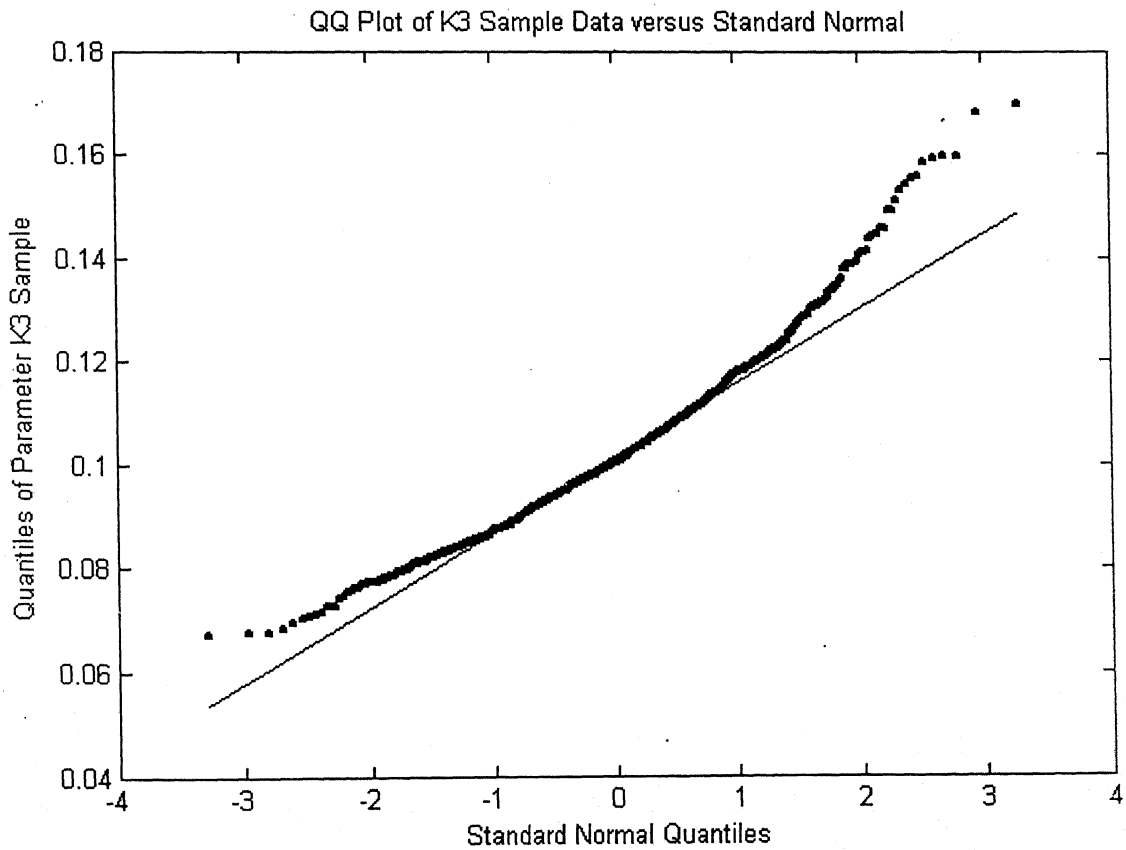
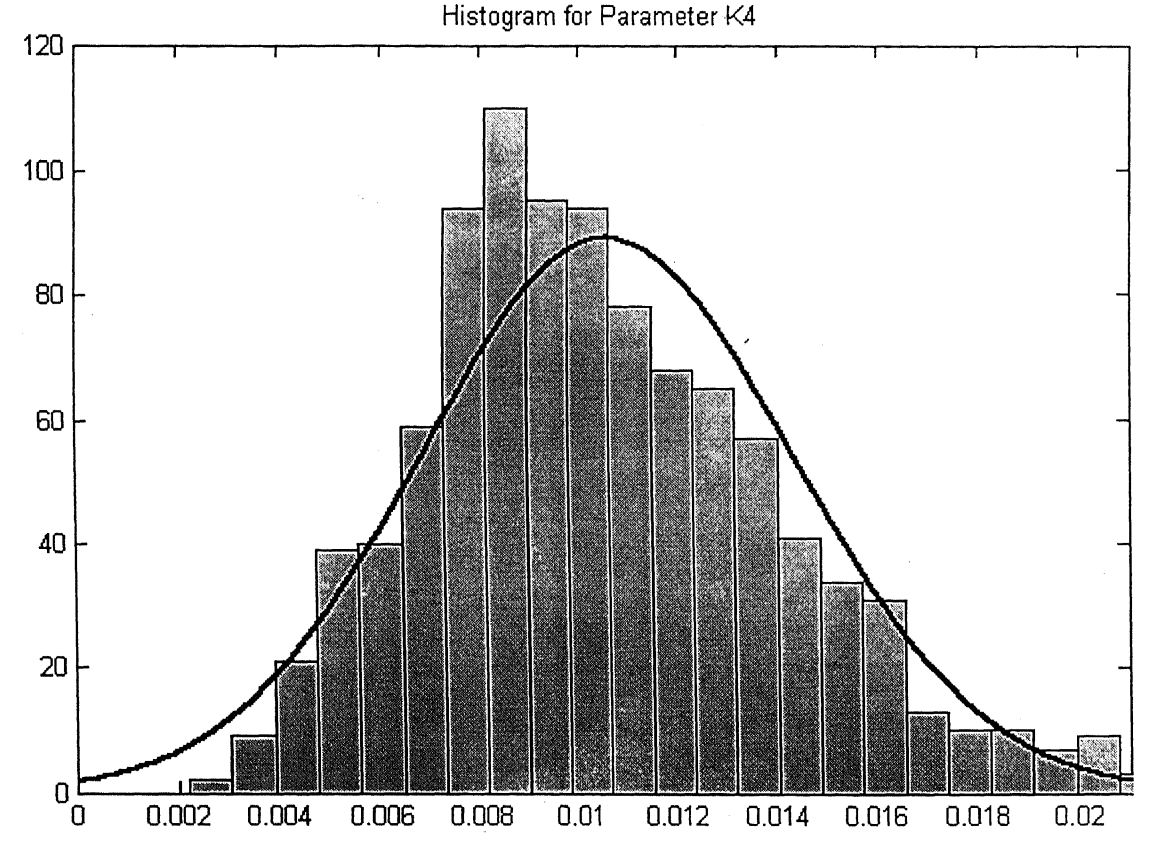


Fig 4.9 Histogram & Q-Q Plot for Parameter K_3 (n=15)



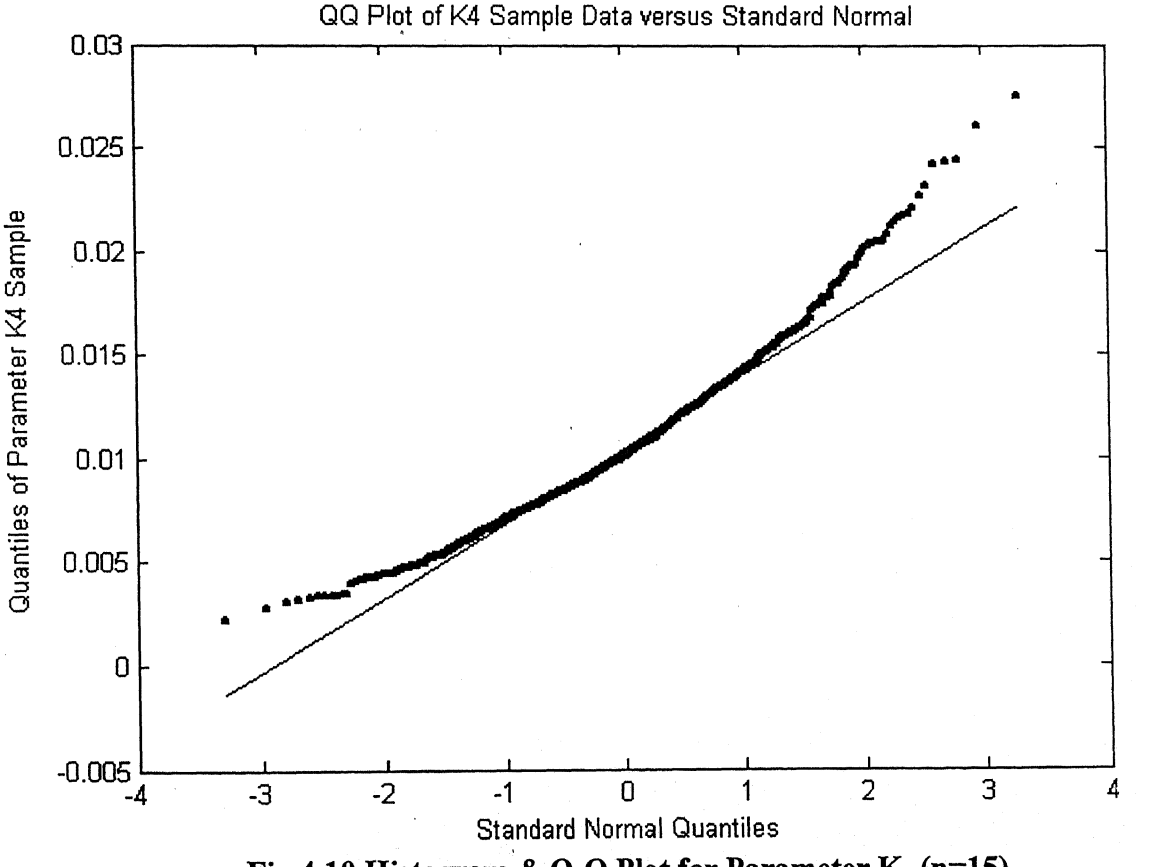
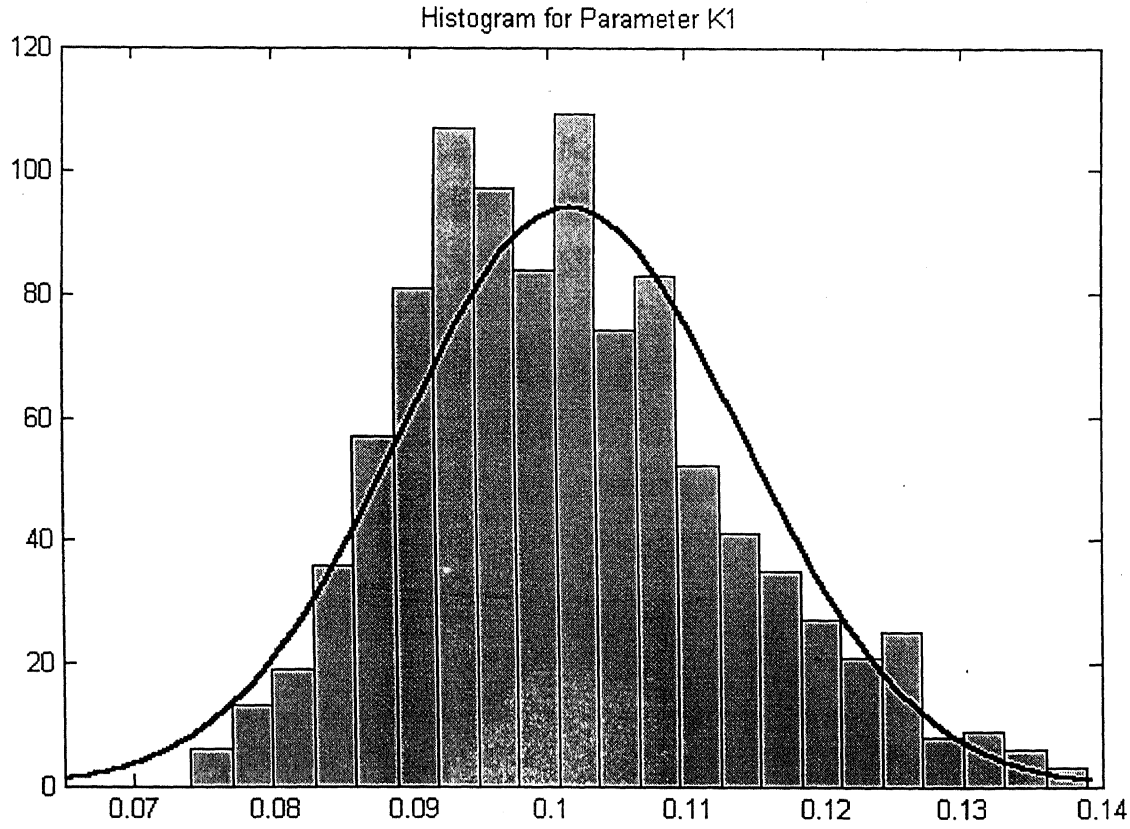


Fig 4.10 Histogram & Q-Q Plot for Parameter K_4 (n=15)



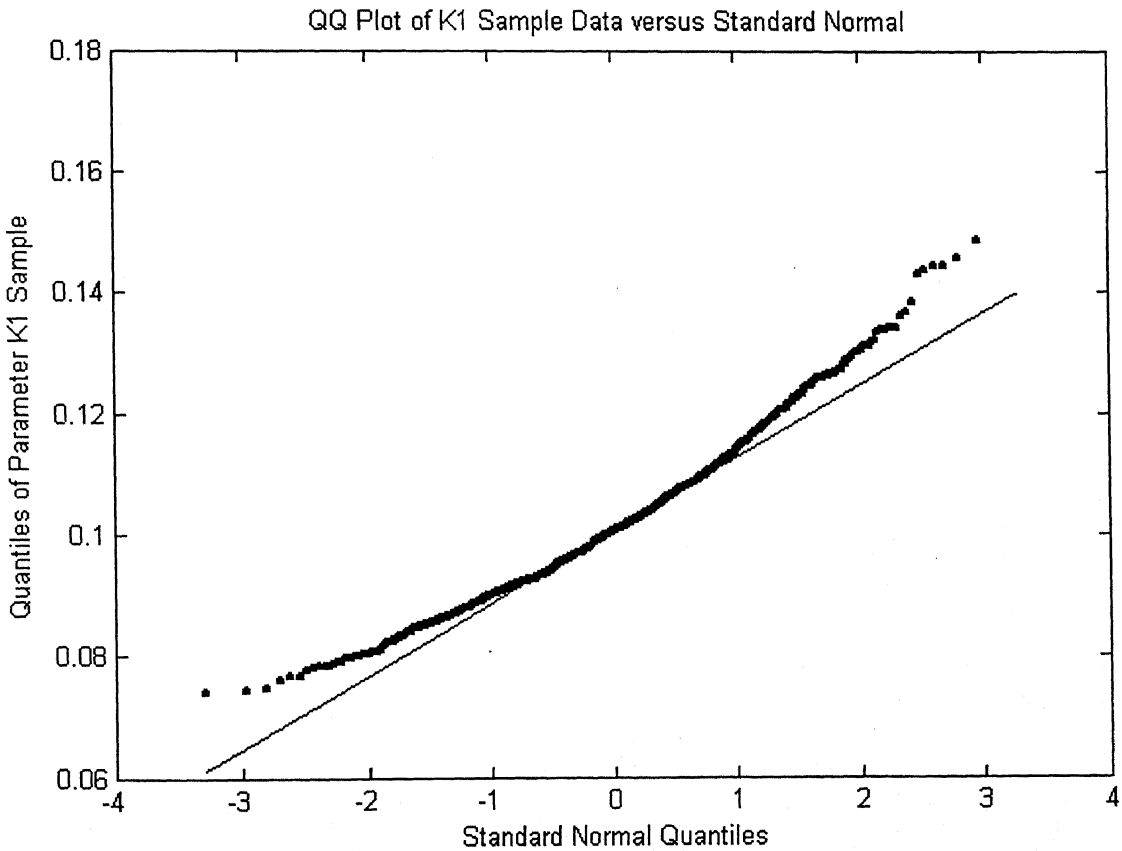
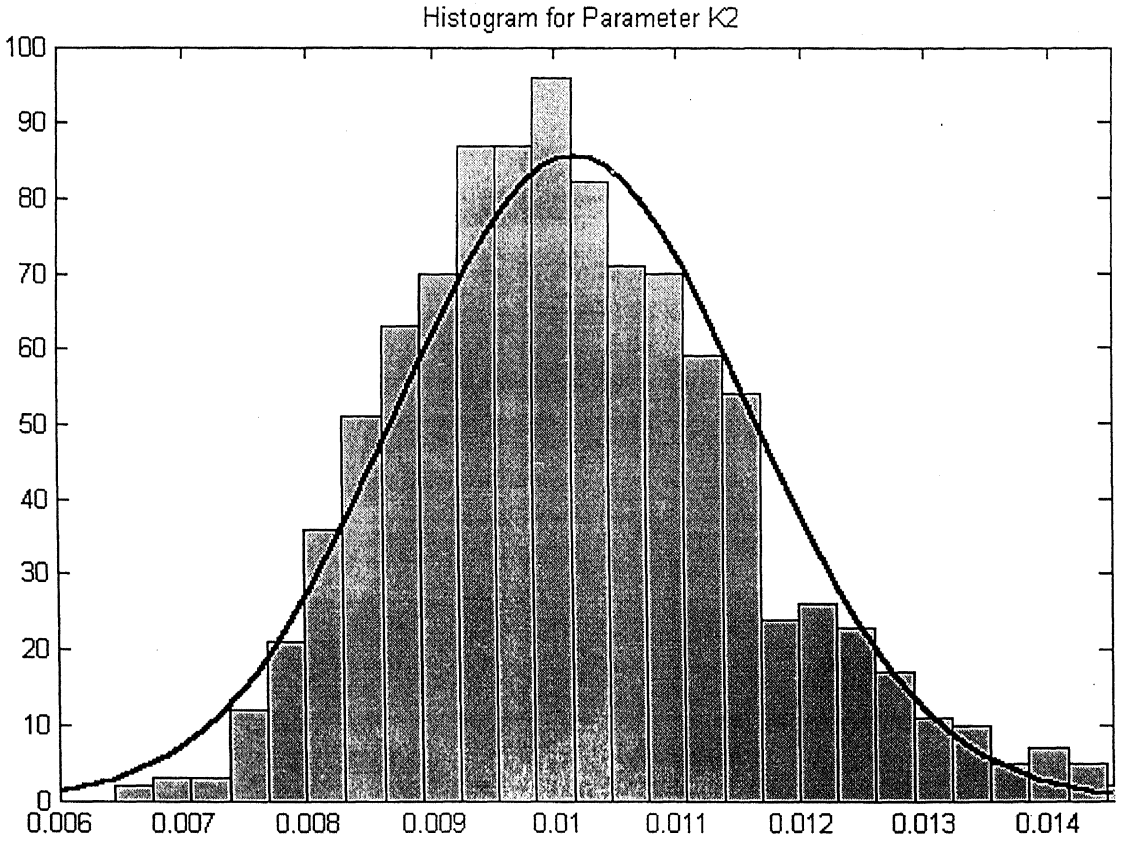


Fig 4.11 Histogram & Q-Q Plot for Parameter K_1 (n=19)



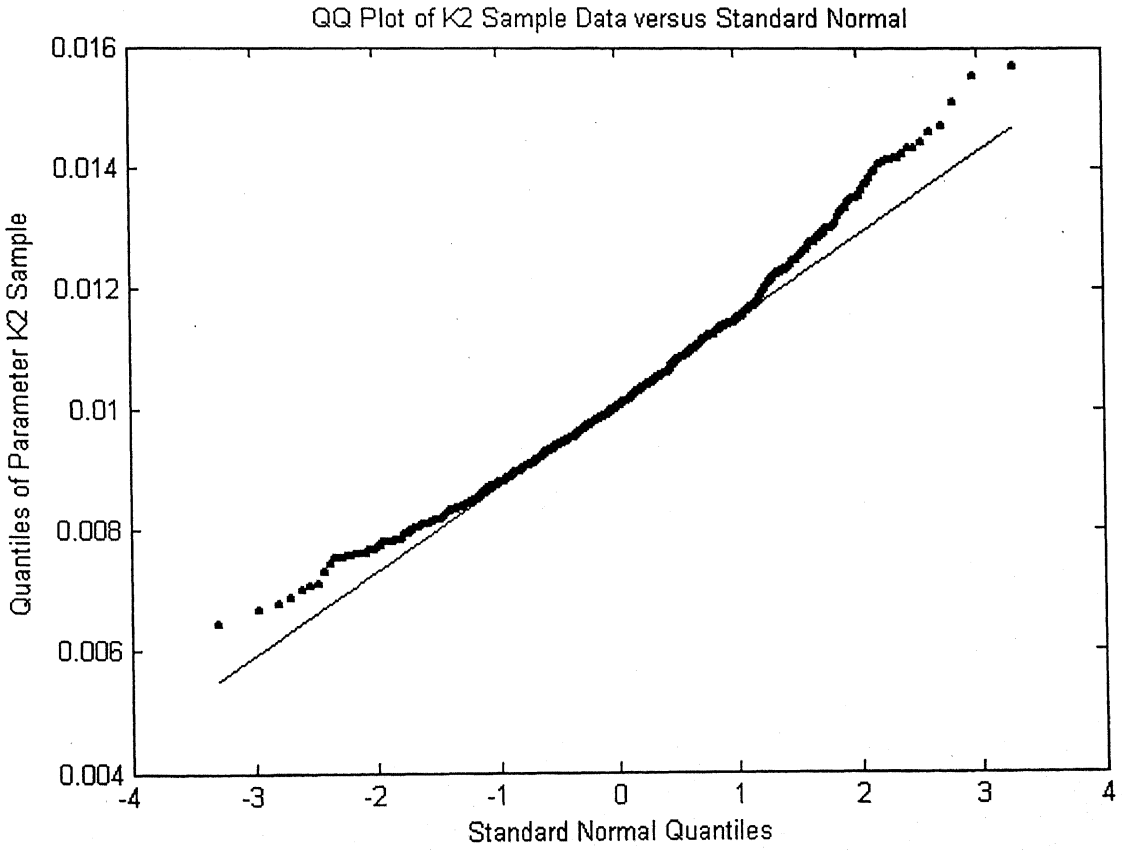
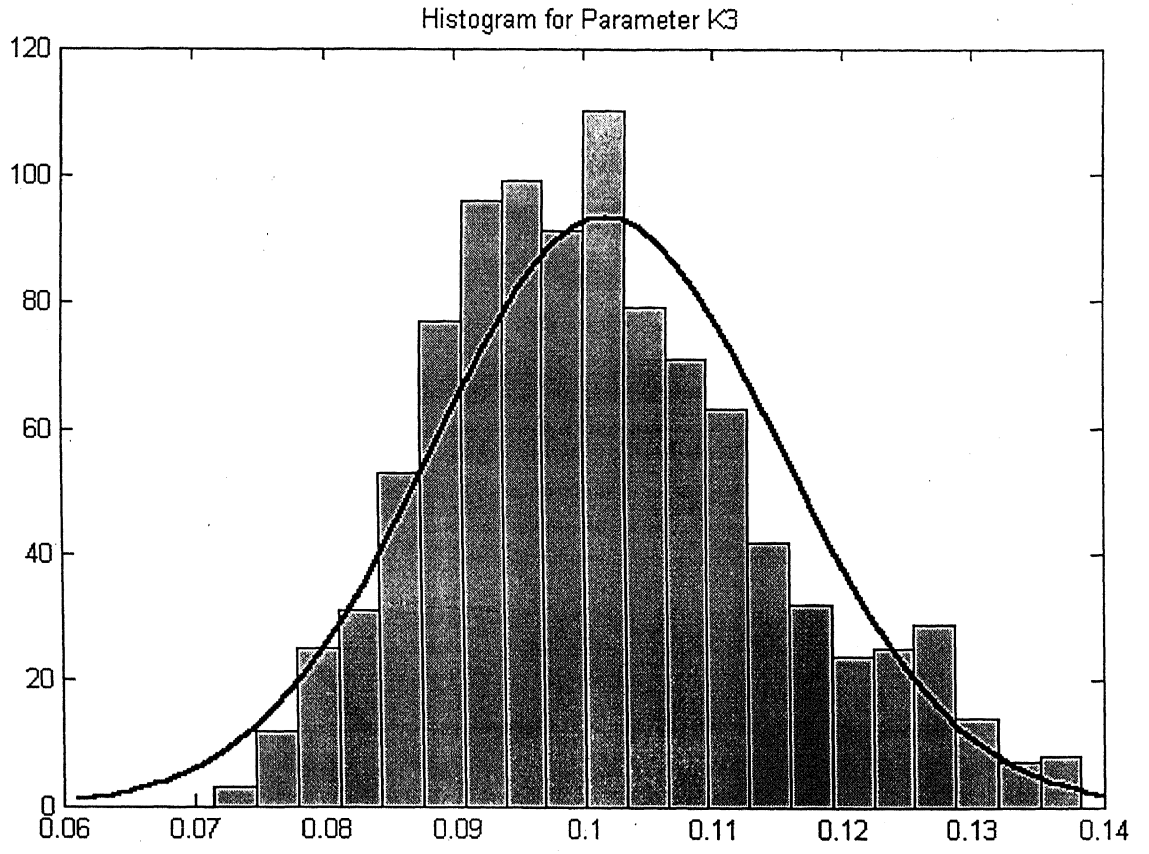


Fig 4.12 Histogram & Q-Q Plot for Parameter K_2 (n=19)



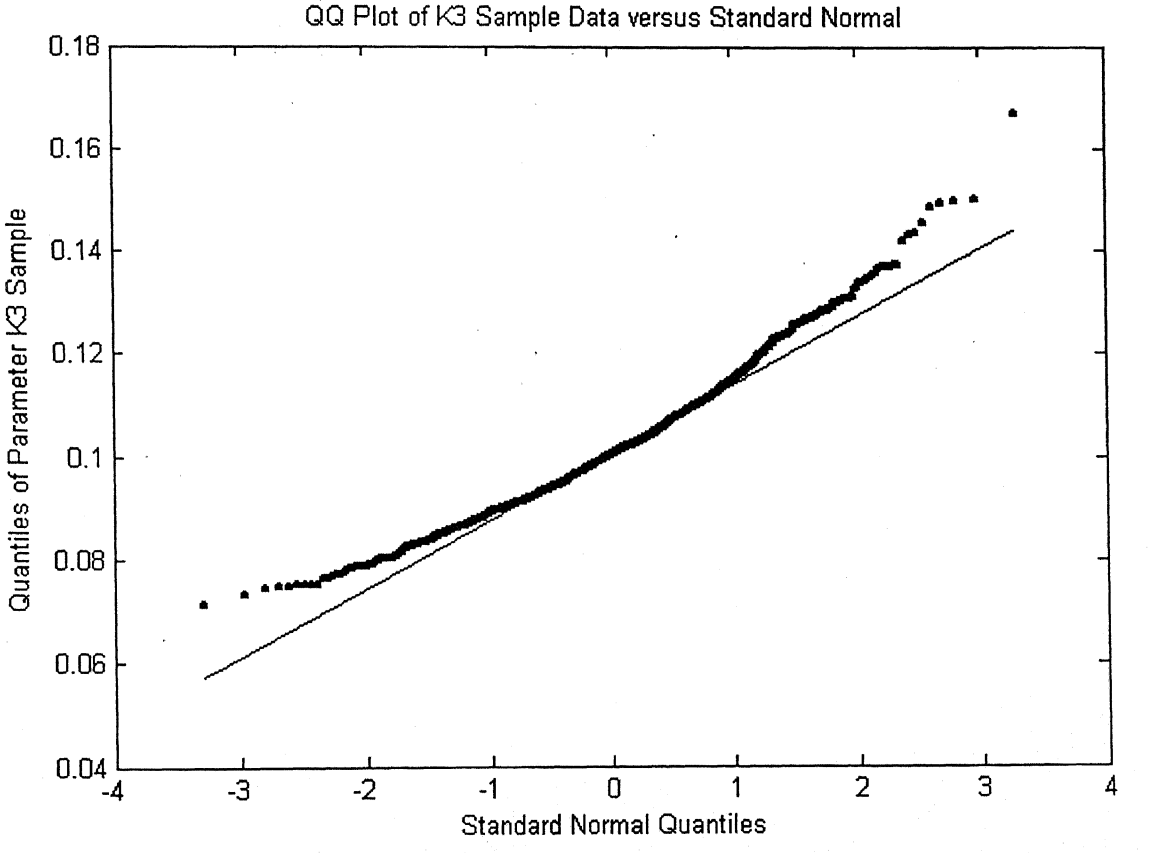


Fig 4.13 Histogram & Q-Q Plot for Parameter K_3 (n=19)

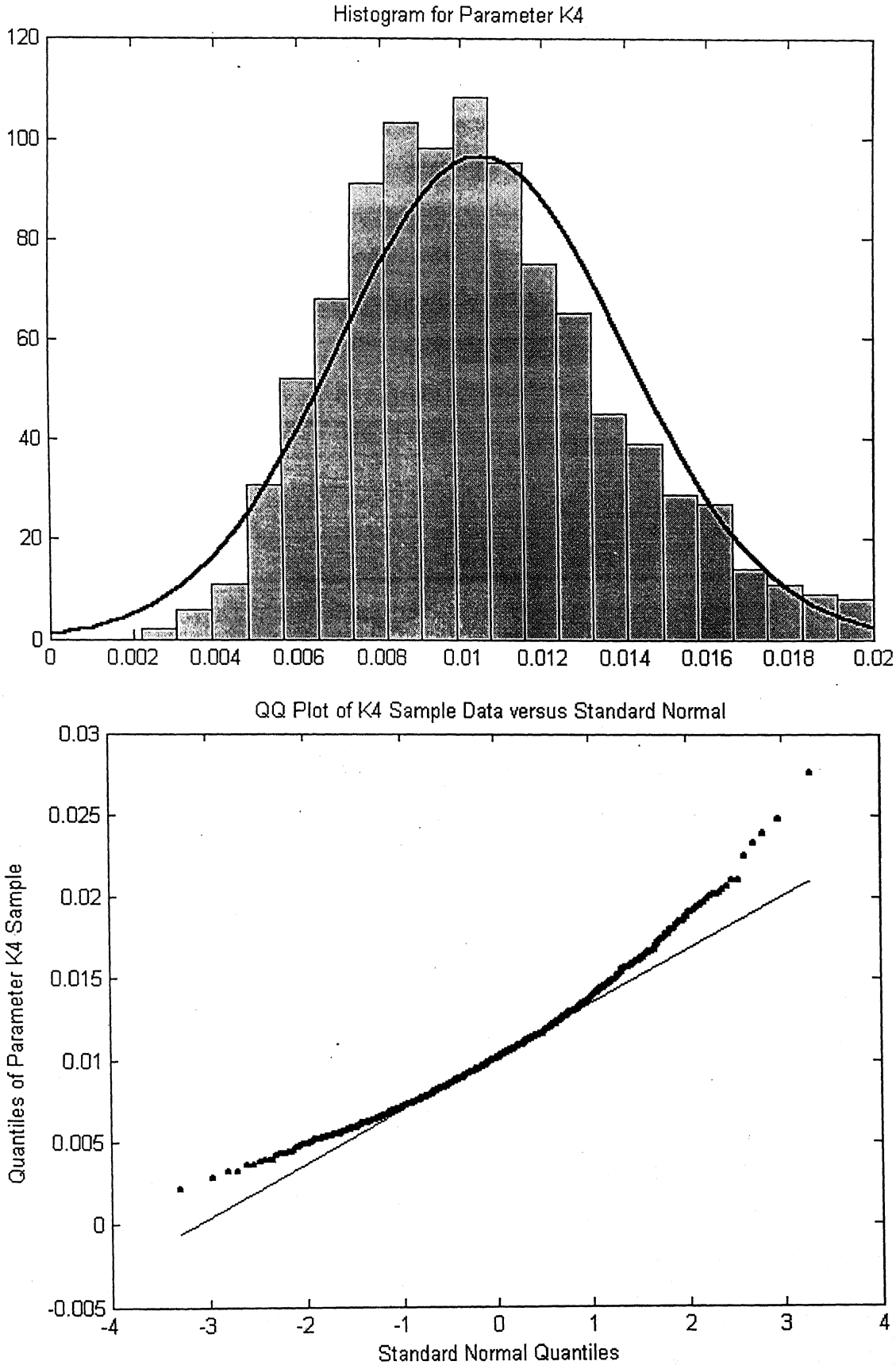


Fig 4.14 Histogram & Q-Q Plot for Parameter K_4 (n=19)

4.2.4 Model Selection

To see whether the model, from which we generated the data turns out to be true, the task of model discrimination has been performed. The cross validation method is used for model discrimination. The estimates of parameters obtained by both the methods of parameter estimation have been used for model selection. In the cross validation method the model, which has the smallest prediction error is considered as the best one. The value of the cross validation index was calculated using the Equation 3.66 & 3.67. The results are shown in the following Table 4.5.

Table 4.14 Cross Validation Results

σ_1^2, σ_2^2	CV(1)	CV(2)	CV(3)	CV(4)
0.5, 0.5	21.5970	25.9055	3554.9	21.6166
0.5, 0.75	22.3985	29.3677	1414.2	22.4418
0.5, 1.0	31.5822	33.6906	1463.3	32.9154
0.5, 2.0	42.3259	45.0432	1501.6	43.9629

From the above results it is explicit that for different values of variances, the cross validation index for model 1 is the least among all the rival models. This shows that model 1 is the best among all the rival models. It is true since we generated the data from model 1 with different variances. The studies made by Buzzi et. al.[10] also says that model 1 is the best one with the help of posterior probabilities obtained by sequential design of experiments.

4.3 Application to the real example

Carr [11] discussed the catalytic isomerization of n-pentane and gave the kinetics for the reaction. Using the data given, the parameters were estimated by taking it as a single response system.

4.3.1 Single response system

The reaction scheme, which is explained by Carr [11] for the isomerization of n-pentane, involves two mechanisms. They are single site mechanism and dual site mechanism, which will be discriminated with the help of parameter estimation. The reaction is given as follows.

$$n-C_5H_{12} \rightleftharpoons i-C_5H_{12} \tag{4.3}$$

For single site mechanism (Model 1) the rate equation will be represented by,

$$y_s = \frac{k_0 k_2 (C_2 - C_3/1.63)}{(1 + k_1 C_1 + k_2 C_2 + k_3 C_3)}. \tag{4.4}$$

For dual site mechanism (Model 2) the rate equation is given by,

$$\grave{y}_d = \frac{k_0 k_2 (C_2 - C_3/1.63)}{(1 + k_1 C_1 + k_2 C_2 + k_3 C_3)^2}. \tag{4.5}$$

In the present work the theory of least square estimation is used for the estimation of the parameters involved in each of the response. The Objective functions are defined as the residual sum of squares. Considering the same kinetics as given in the literature the two objective functions are defined as

$$(O.F.)_1 = \underset{k_0,k_1,k_2,k_3}{Min} \sum \left(\frac{k_0 k_2 (C_2 - C_3/1.63)}{(1 + k_1 C_1 + k_2 C_2 + k_3 C_3)} - y_s \right)^2, \tag{4.6}$$

$$(O.F.)_2 = \underset{k_0,k_1,k_2,k_3}{Min} \sum \left(\frac{k_0 k_2 (C_2 - C_3/1.63)}{(1 + k_1 C_1 + k_2 C_2 + k_3 C_3)^2} - y_d \right)^2. \tag{4.7}$$

4.3.2 Parameter Estimation

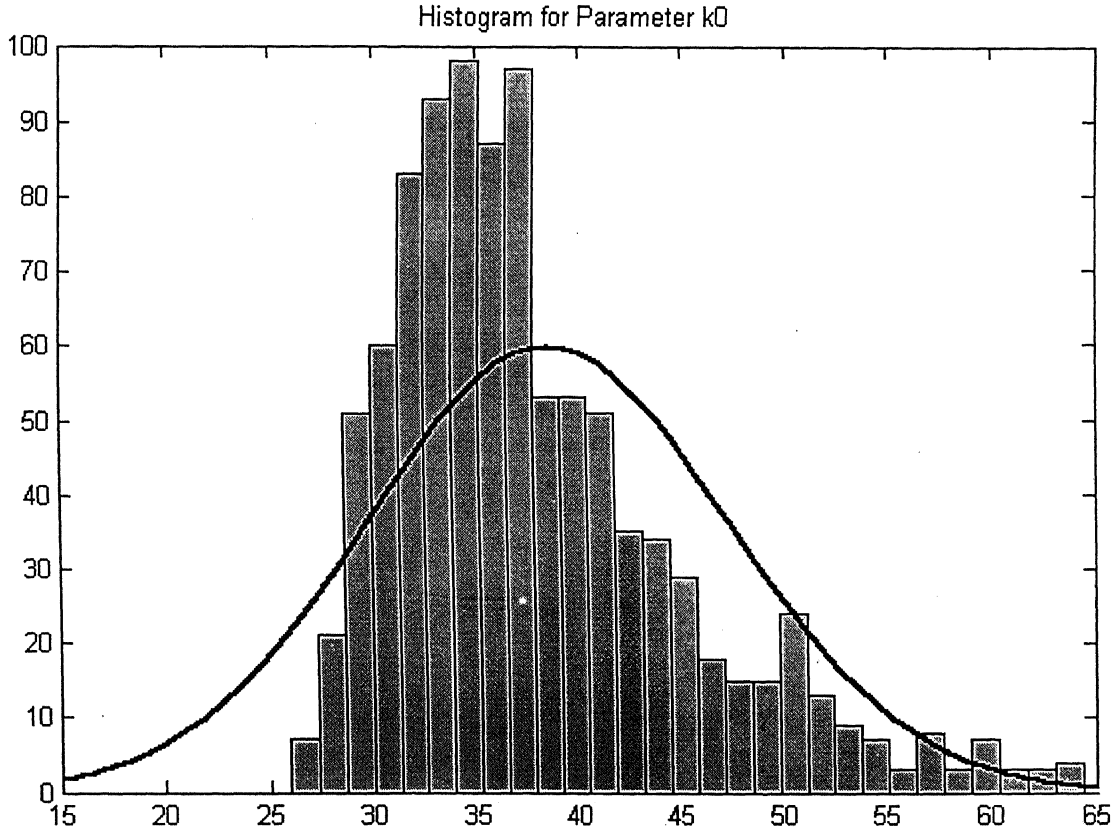
The best values of the parameters were calculated using a non-linear regression algorithm based upon the function minimization technique of GNLM. The parameters obtained are compared with the values given by Jhonson et. al.[29] as follows .

Table 4.15 Comparison of the Estimated Parameters

Parameters	Model 1		Model 2	
	Present Study	Jhonson et. al.[29]	Present Study	Jhonson et. al.[29]
k_0	35.9071	36.21	133.0124	132.97
k_1	0.0709	0.9	0.0022	0.033
k_2	0.0378	0.48	0.0013	0.019
k_3	0.1671	2.13	0.0049	0.072
SOS	3.235	4.31	3.478	4.88

The variance of the above models is calculated by using,

$$\hat{\sigma}^2 = \frac{SOS}{d.f.}. \tag{4.8}$$



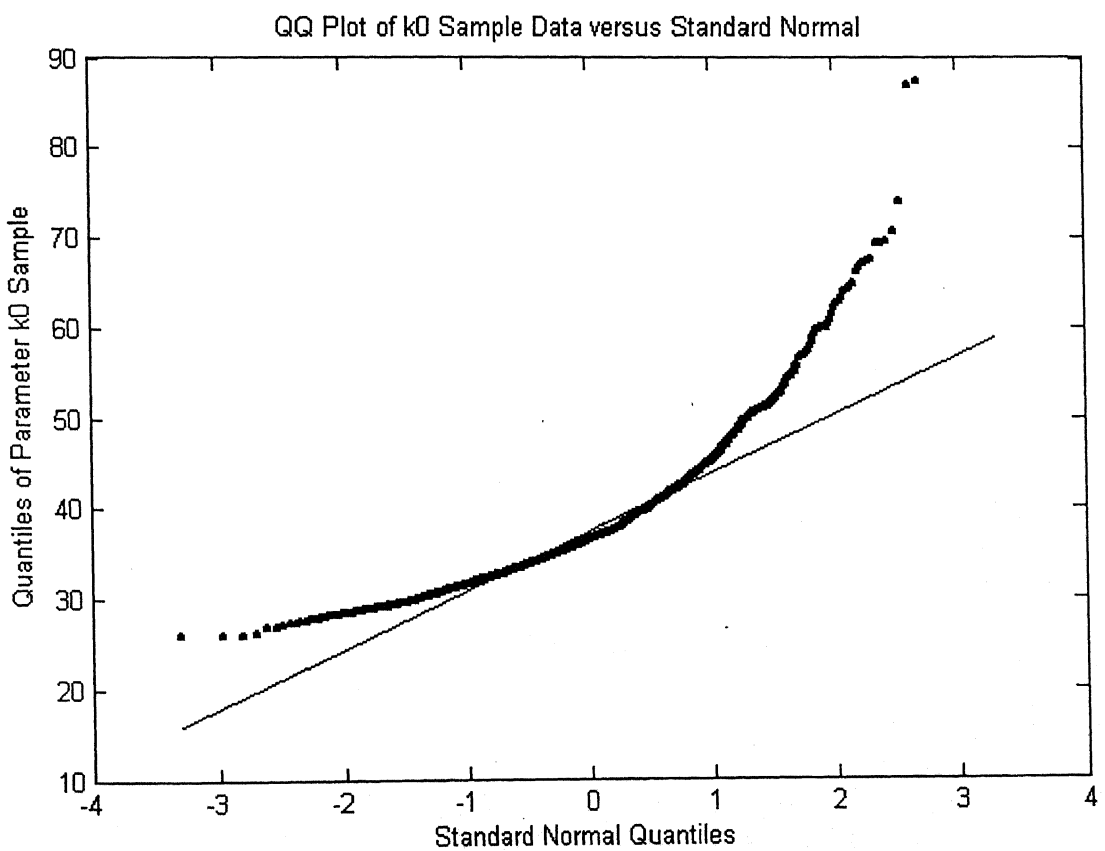
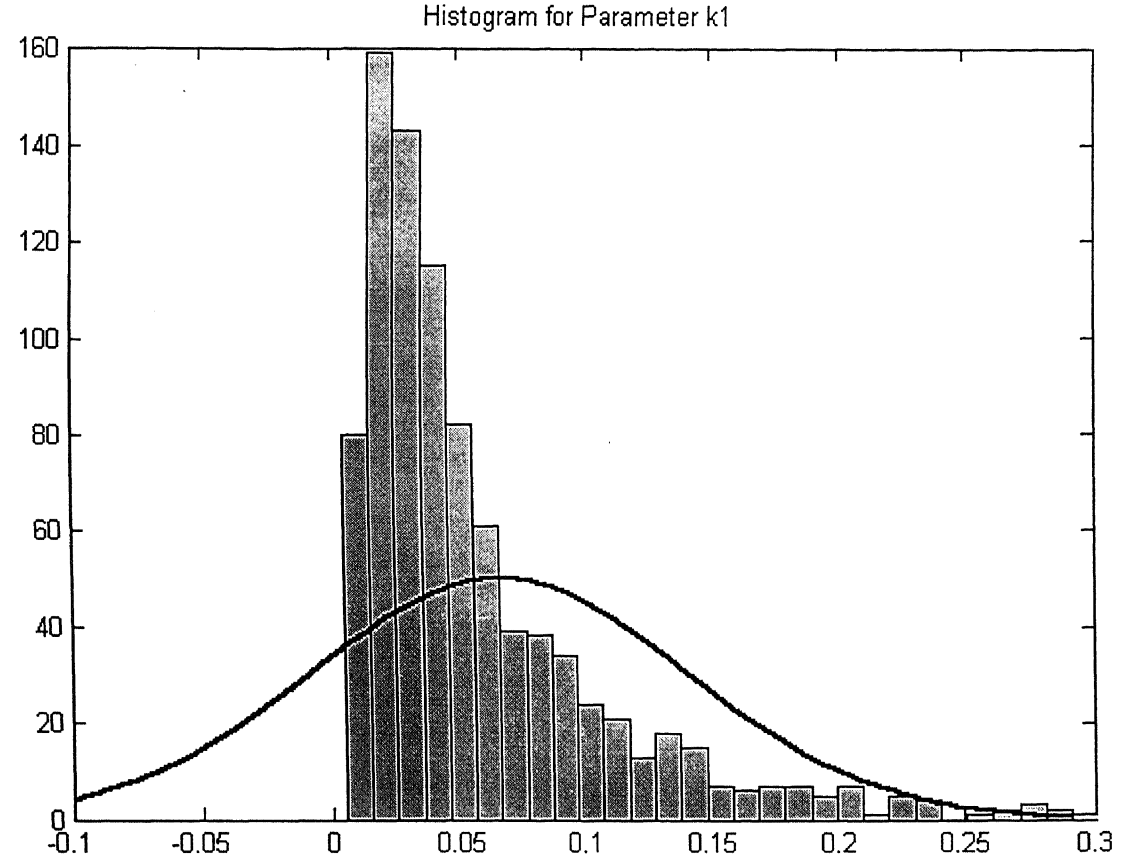


Fig. 4.15 Histogram and Q-Q Plot for Parameter k_0



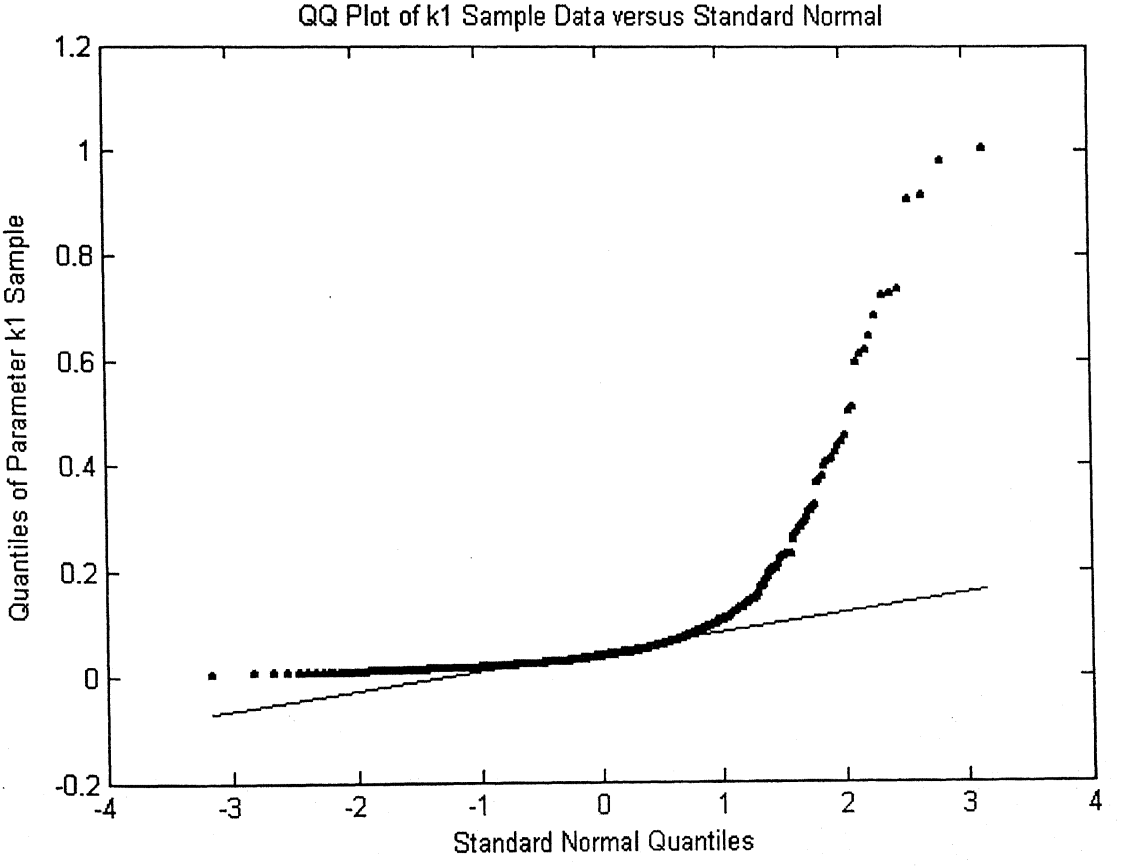


Fig. 4.16 Histogram and Q-Q Plot for Parameter k_1

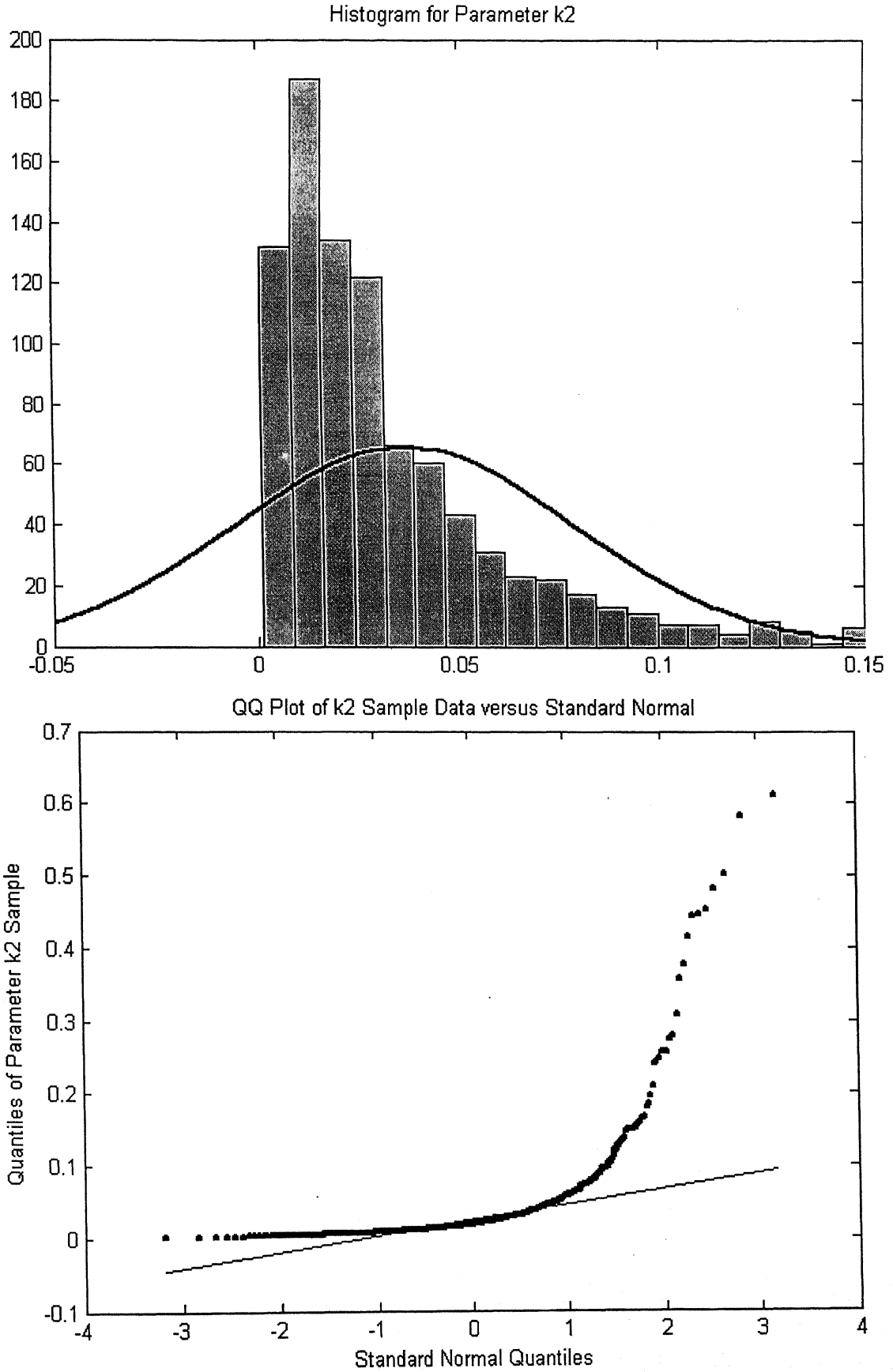


Fig. 4.17 Histogram and Q-Q Plot for Parameter k_2

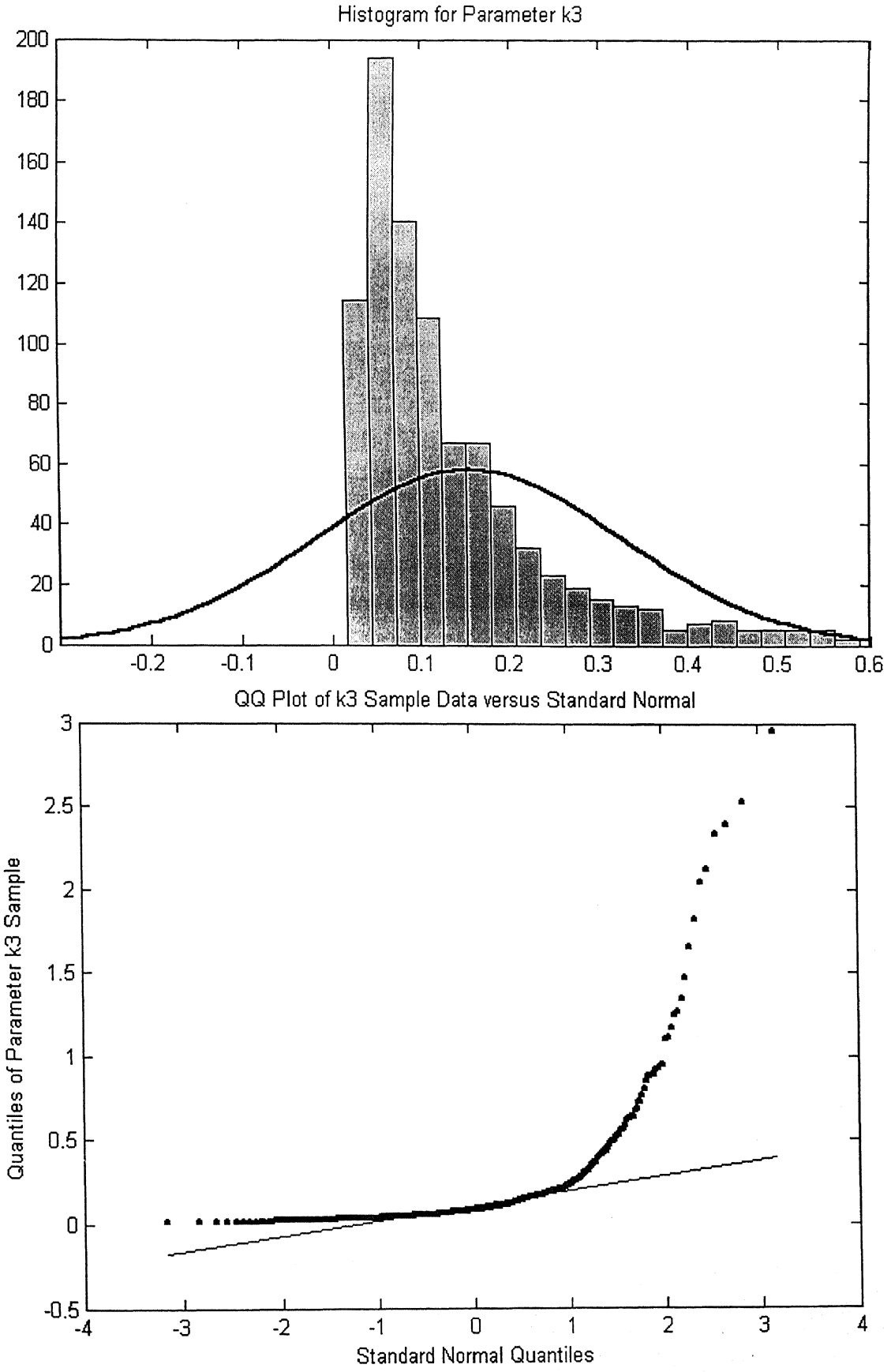
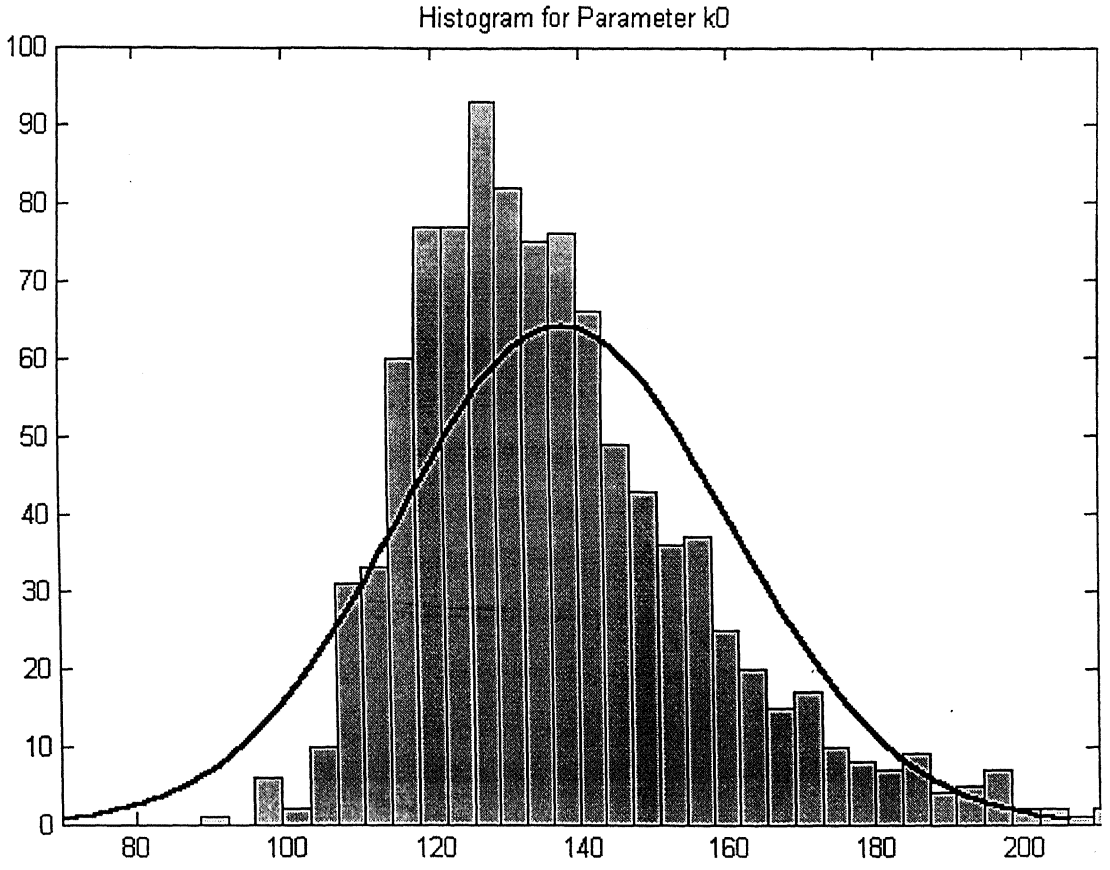


Fig. 4.18 Histogram and Q-Q Plot for Parameter k_3



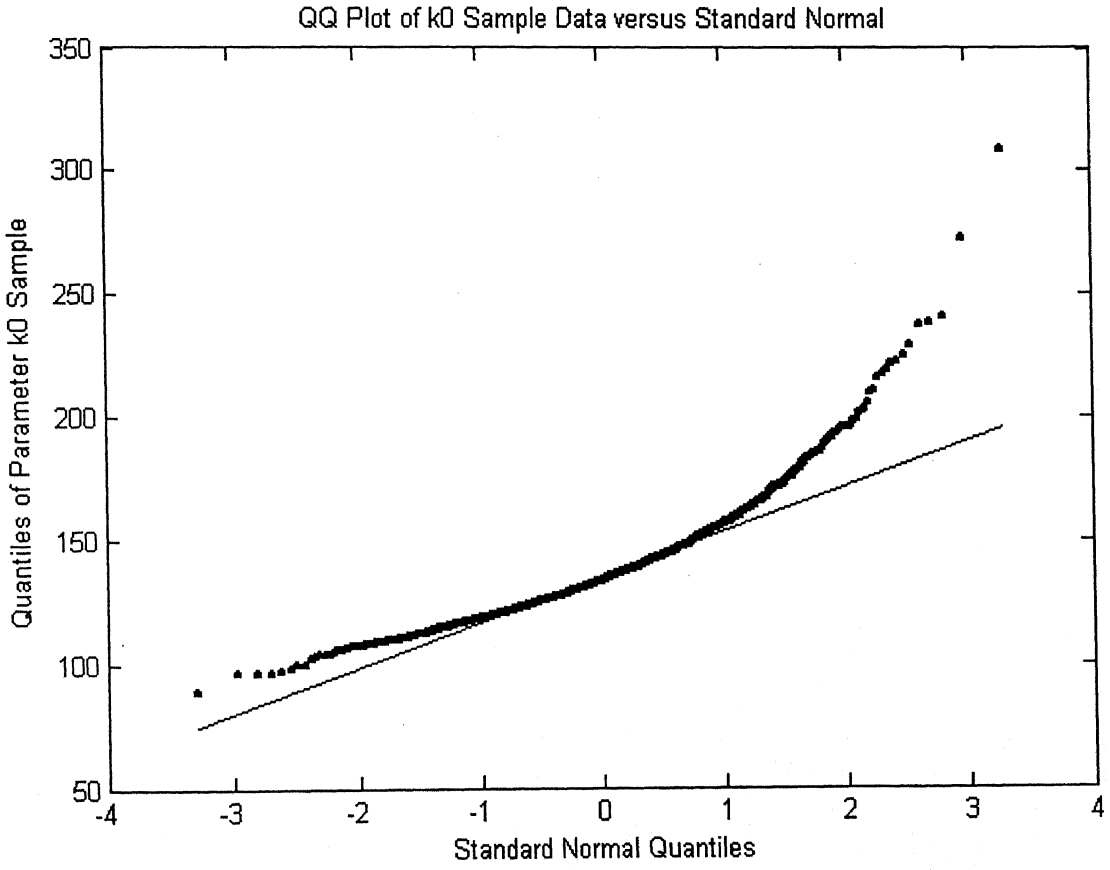
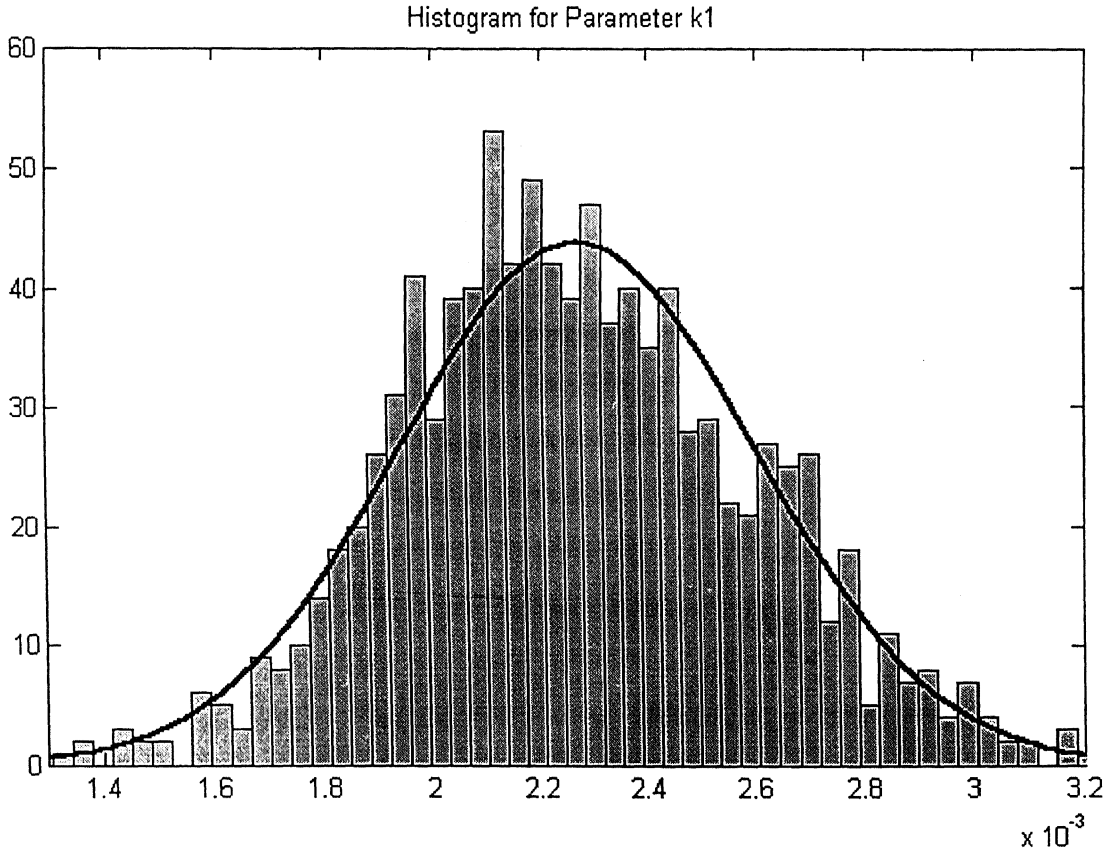


Fig. 4.19 Histogram and Q-Q Plot for Parameter k_0



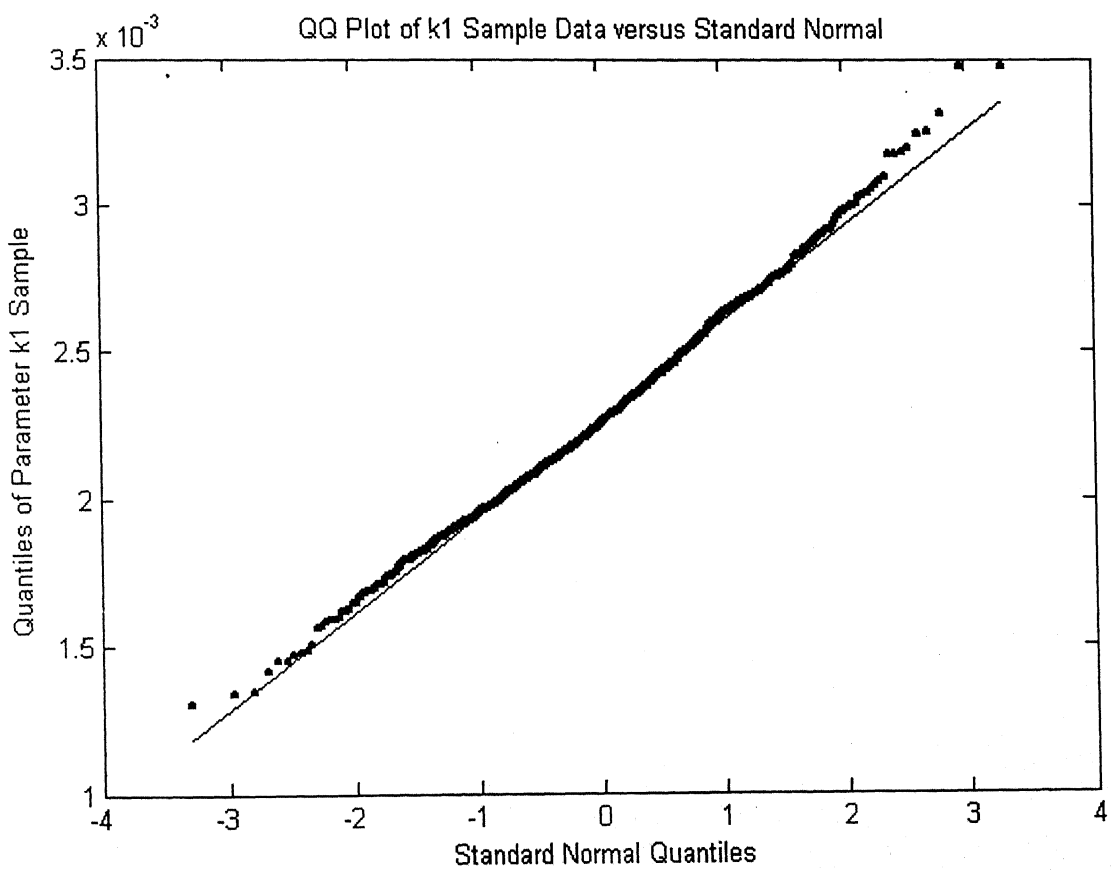
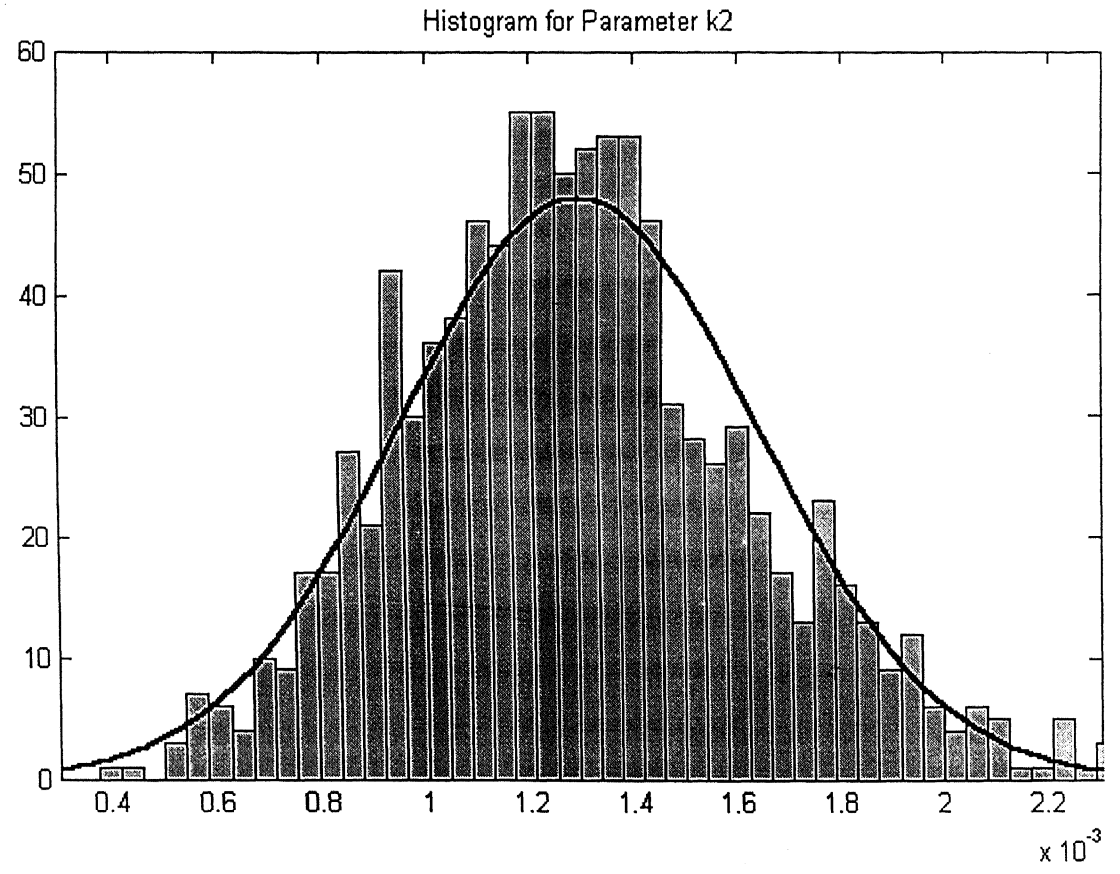


Fig. 4.20 Histogram and Q-Q Plot for Parameter k_1



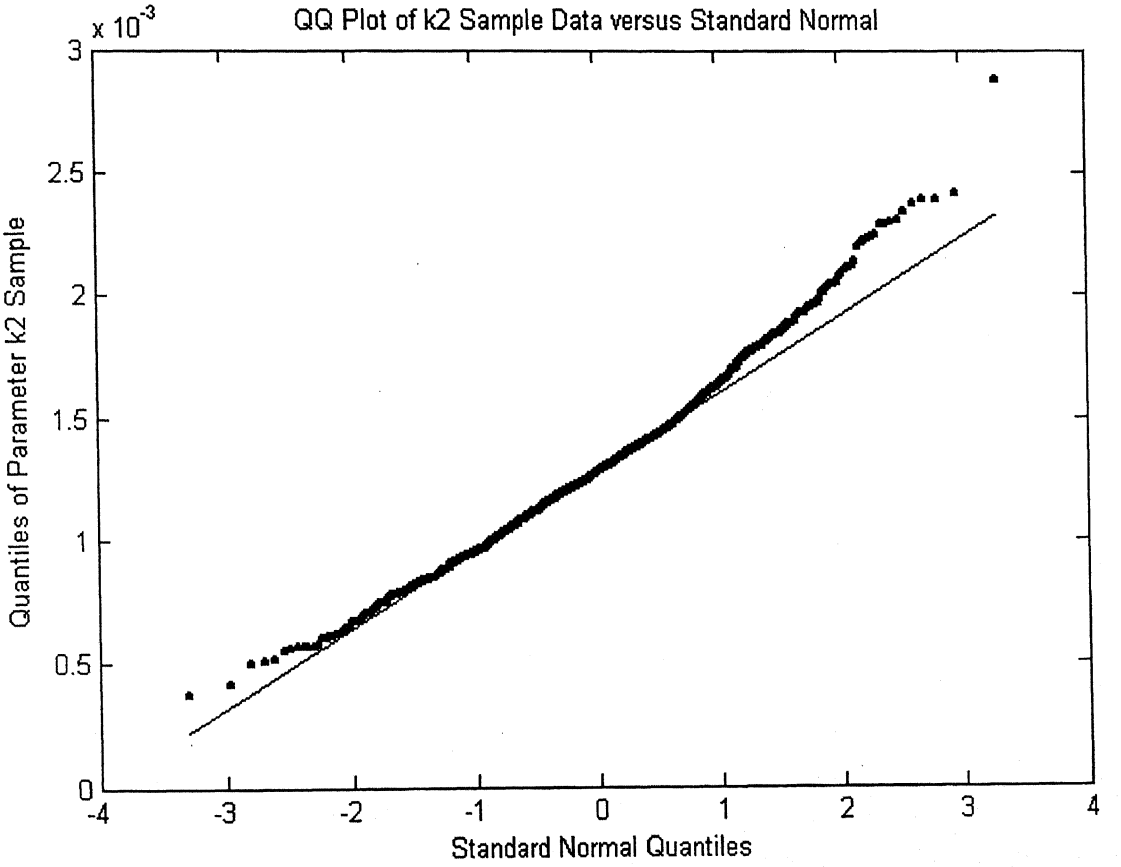
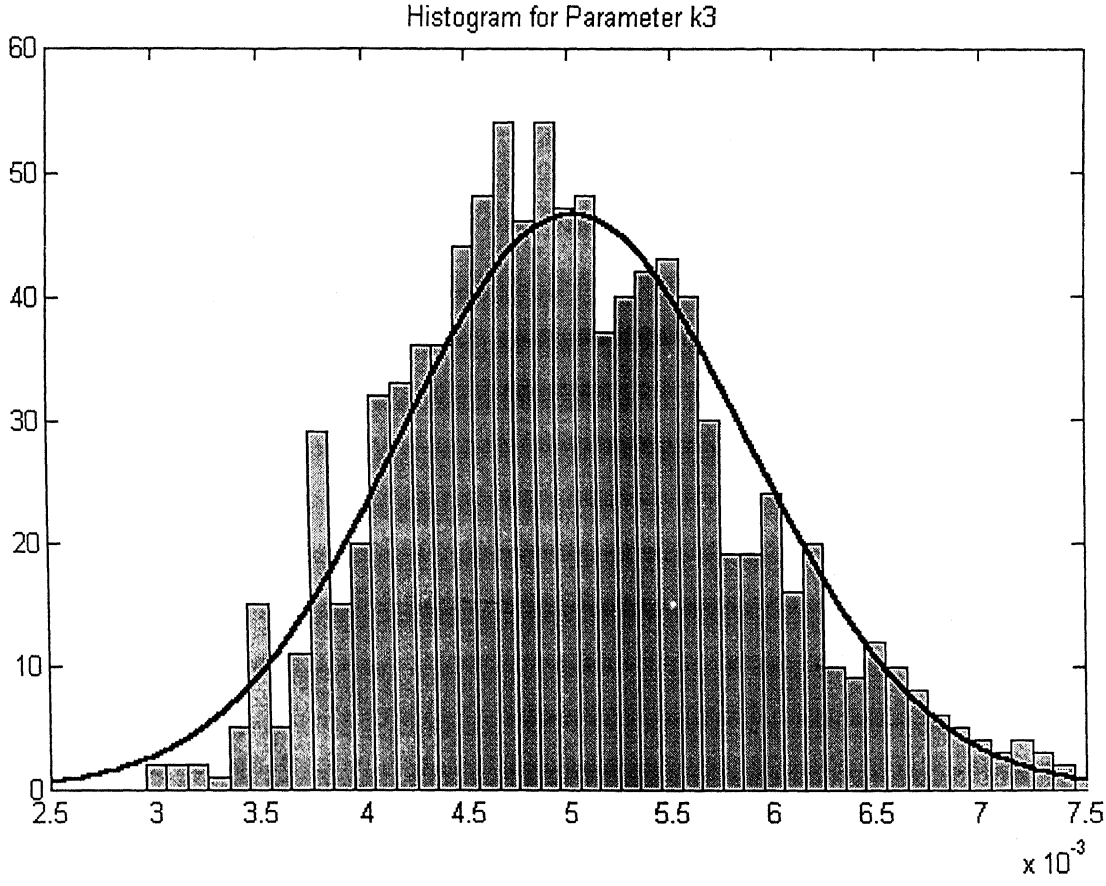


Fig. 4.21 Histogram and Q-Q Plot for Parameter k_2



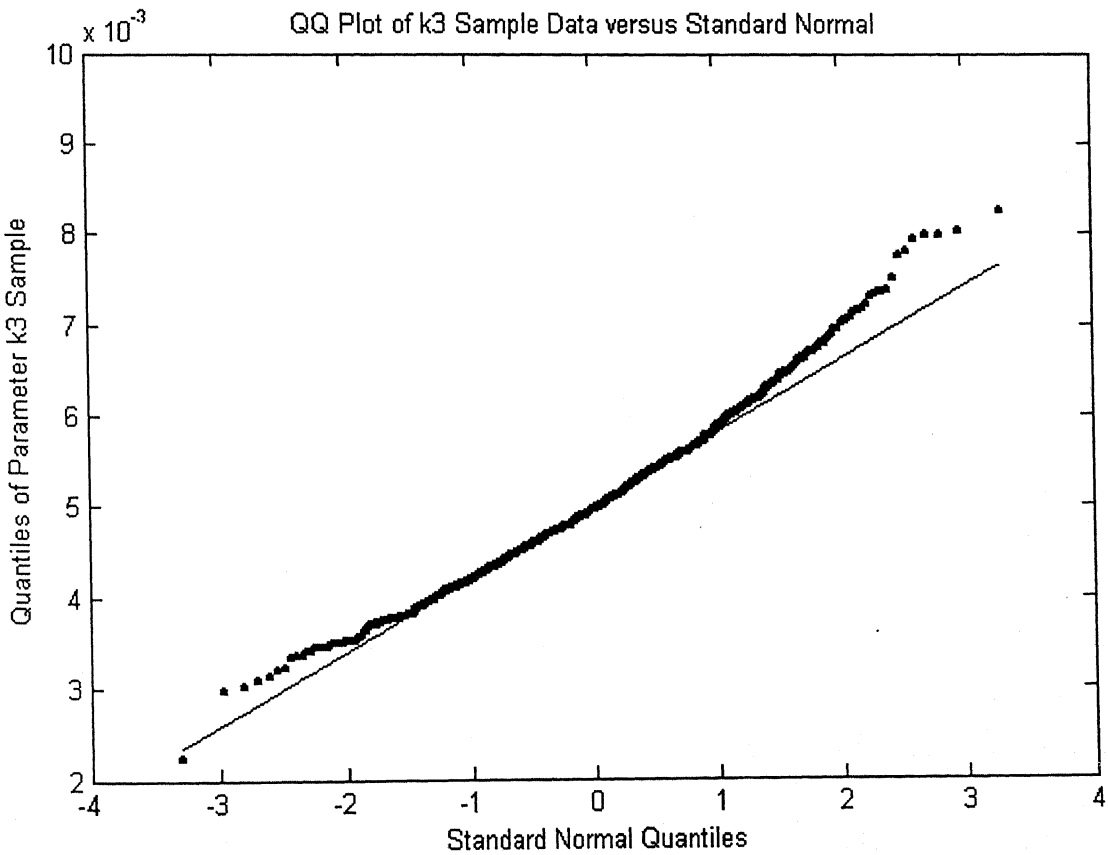


Fig. 4.22 Histogram and Q-Q Plot for Parameter k_3

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In this dissertation some aspects of parameter estimation and model selection have been studied for single response and multi response systems, first with the simulated data and then with real life examples. The GNLM method is used for minimization of the residual sum of squares based on Asymptotic theory and Bootstrapping Methods. The program converges to minima that depend on the initial guesses, which should be taken properly in order to get the global minima. The case study of the single response system has been undertaken with a view to comparing the estimates of parameters and residual sum of squares with those reported in the literature.

Based on this study the following conclusions can be drawn.

1. In order to obtain the global minima, different initial guesses are used in the GNLM method and they are found to give same estimates.
2. The efficiency of Least Square Estimates (LSE) is tested with the help of asymptotic theory and found that the assumptions of this theory are satisfied for large sample sizes $(n \geq 15)$.
3. The average estimates and their mean squared errors give an insight into the robustness of the parameters.
4. The asymptotic confidence intervals and coverage probabilities give an assurance to the 95% confidence interval of the parameters.
5. More over bootstrapping confidence intervals and coverage probabilities works quite well even for lower sample sizes.
6. From the normality plots of parameters obtained by simulation studies with different variances says the parameters are unbiased and normally distributed.
7. Cross validation procedure can be used as a model selection technique in competing situations.

5.2 Recommendations

In the minimization of the objective function, the initial guesses seem to be very important. Appropriate initial guesses lead to the identification of the global minima. The estimation of the parameters first with some simpler methods (for example by transferring the model to a linear one if possible) is recommended as these estimates of parameters can be used as the initial guesses for the methods in which the initial guesses are required.

It is also recommended that one can use evolutionary algorithms like GA and SA, which does not require initial guesses but bounds of variables. The disadvantage of these methods is they are very time consuming. The combination of Asymptotic theory and Bootstrapping methods with Evolutionary algorithms can be tried to study the behavior of parameter estimation in nonlinear regression.

Study of parametric and non-parametric bootstrapping methods on parameter estimation is recommended as the future work.

BIBLIOGRAPHY

1. Bard Y. and Lapidus L., Kinetics analysis by digital parameter estimation, *Catalysis Rev.*, **1968**, 2, 67-73.
2. Bard Y., Nonlinear Parameter Estimation, *Academic Press*, New York **1974**.
3. Bates D.M. and Watts D.G., Nonlinear Regression Analysis and Its Applications, *Wiley*, New York **1998**.
4. Bickel and Freedman, Asymptotic normality and the bootstrap in stratified sampling, *Ann. Statist.*, **1984**, 12, 470-482.
5. Boag I.F., Bacon D.W. and Downie J., Using a statistical Multi response method of experimental design in a reaction network study, *Can. J. Chem. Eng.*, **1978**, 56, 389-395.
6. Box G. E. P. and Draper N. R., The Bayesian estimation of common parameters from several responses, *Biometrika*, **1965**, 52, 3, 355-365.
7. Box G.E.P., Hunter W.G. and Hunter J.S., Statisitics for Experimenters, *John Wiley and Sons*, **1978**.
8. Briemn, The little bootstrap and other methods for dimensionality selection in regression, *J. Am. Statist. Assoc.*, **1992**, 87,738-754.
9. Buzzi F. G. and Forzatti P., A new sequential experimental design procedure for discriminating among rival models, *Chem. Eng. Sci.*, **1983**, 38, 225-232.
10. Buzzi F. G., Forzatti P., Emig G. and Hofmann H., Sequential experimental design for model discrimination in the case of multiple responses, *Chem. Eng. Sci.*, **1984**, 39, 81-85.
11. Carr N.L., Kinetics of catalytic isomerization of n-Pentane, *Ind. Eng. Chem.*, **1960**, 52, 391-399.
12. Chernick M.R., Bootstrap Methods, A Practitioner's guide, *Willey* series in *Prob. & Stat.*, **1999**.
13. Cox D.R., Tests of separate families of hypotheses, *Proc. Fourth Berkeley Sym.*, **1961**, 1, 105.
14. Cox D.R., Further results on tests of separate families of hypotheses, *J. Roy. Stat. Soc., Series B*, **1962**, 24, 406.

15. Draper N. R. and Hunter G. W., Design of experiments for parameter estimation in multi response situations, *Biometrika*, **1961**,53, 3, 525.

16. a) Draper N. R. and Smith H., Applied Regression Analysis, *John Wiley & Sons, Inc.*, Second Edition, **1981**.

    b) Draper N. R. and Smith H., Applied Regression Analysis, *John Wiley & Sons, Inc.*, Third Edition, **1998**.

17. Efron, The Jackknife, The Bootstrap and the other resampling plans, *SIAM*, **1982a**.

18. Efron, Bootstrap confidence interval for a class of parametric problems, *Biometrika*, **1985**, 72, 45-58.

19. Efron, Better bootstrap confidence intervals, *J. Am. Statist. Asoc.*, **1987**, 82,171-200.

20. Efron B. and Tibishirani R., Bootstrap methods for standard errors: Confidence interval and measures of statistical accuracy, *Statist. Sci.*, **1986**, 1, 54-77.

21. Efron B. and Tibishirani R., An Introduction to the Bootstrap, *Chapman & Hall*, **1993**.

22. Freedman, Bootstrapping regression models, *Ann. Statist.*, **1981**, 9, 1218-1228.

23. Freedman and Peters, Bootstrapping a regression equation: Some empirical results, *J. Am. Statist. Assoc.*, **1984a**, 79, 97-106.

24. Froment G. F. and Bischoff K. B., Chemical Reactor Analysis and Design, *John Wiley and Sons*, Second edition, **1979**.

25. Hayes and Perl, Application of bootstrap statistical method to tau-decay-mode problem, *Phys. Rev. D*, **1989**, 39, 274-279.

26. Hill W. J., Statistical techniques for model building, *PhD. Thesis, University of Wisconsin*, **1966**.

27. Hocking R.R., The analysis and selection of variables in linear regression, *Biometrics*, **1976**, 32, 1-49.

28. Hunter W.G. and Reiner A.M., Designs for discriminating between two rival models, *Technometrics*, **1965**, 7, 307.

29. Jhonson R.A., Standal N.A. and Mezaki R., Re-Examination of Hougen-Watson rate model including a weighted least squares approach, *I & EC Fundamentals*, **1968**, 7, 181-184.

30. Kanodia R., Parameter estimation and model discrimination in single response models, *M. Tech. Thesis, I.I.T. Kanpur*, **2001**.
31. Kundu D. and Murali G., Model selection in linear regression, *Comp. Statist. & Data Anal.*, **1995**, 22, 461-469.
32. Lehmann E.L. and Casella G., Theory of Point Estimation, *Springer-Verlog*, New York, Second Edition, **1998**.
33. Mezaki R. and Butt J. B., Estimation of rate constants from multi response kinetic data, *Ind. Eng. Chem. Fund.*, **1968**, 7 , 120.
34. Moros R., Kalis H., Rex H.G. and Schaffarczyk St., A genetic algorithm for generating initial parameter estimations for kinetics of catalytic processes, *Comp. & Chem. Engg.*, **1996**, 20, 1257-1270.
35. Negi A., Parameter estimation and model discrimination in single and multi response models, *M. Tech. Thesis, I.I.T. Kanpur*, **2003**.
36. Prasad K.B.S. and Rao M.S., Use of expected likelihood in sequential model discrimination in multi response systems, *Chem. Engg. Sc.*, **1976**, 32, 1411-1418.
37. Rohatgi V.K., Statistical Inference, *Wiley Series*, New York, **1984**.
38. Seber G. A. F. and Wild, Non-Linear Regression Analysis, *John Wiley and Sons, Inc.*, **1977**.
39. Seinfeld J.H., Nonlinear estimation theory, *Ind. Engg. Chem.*, **1970**, 62, 32-41.
40. Sen and Srivastava M., Regression Analysis Theory, Methods and Applications, *Springer-Verlag*, New York, **1990**.
41. Shao J., On resampling methods for variance and bias estimation in linear models, *Ann. Statist.*, **1988a**, 16, 986-1008.
42. Shao J., Bootstrap variance and bias estimation in linear models, *Can. J. Statist.*, **1988b**, 16, 371-382.
43. Shao J. and Tu. D., The Jackknife and Bootstrap, *Springer-Verlag*, New york, **1995**.
44. Shimbakuro F.I., Lazer S., Dysen H.B. and Chernick M.R., A quasi-optical method for measuring the complex permittivity of materials, *IEEE Trans., Microwave Theory Tech.*, **1984**, 32, 661-665.

45. Singh S., Design and analysis of experiments for model discrimination in uniresponse and multiresponse systems, *Ph.D. Thesis, I.I.T. Kanpur*, **1986**.

46. Stein K., Fitting kinetic equations to rate data assuming more than one set of active centers on a catalyst, *I & EC Fundamentals*, **1967**, 6,2,169-174.

47. Thompson M.L., Selection of variables in multiple regression, part I & II, *Int. Statist. Rev.*, **1978**, 1-19, 126-146.

48. Vajda S. and Valko Y.P., A direct-indirect procedure for estimation of kinetic parameters, *Comp. & Chem. Engg.*, **1986**, 10, 49-58.

49. Weber, On resampling techniques for regression models, *Statist. Probab. Lett.*, **1984**, 2, 275-278.

50. Wu, Jackknife, Bootstrap and Other resampling plans in regression analysis, *Ann. Statist.*, **1986**, 14, 1261-1350.

APPENDIX I

The simulated multi response data for different variances in the absence of co-variance is given in the following table.

	S.No.	$\sigma_1^2=0$, $\sigma_2^2=0.$	$\sigma_1^2=0.5$, $\sigma_2^2=0.5.$	$\sigma_1^2=0.5$, $\sigma_2^2=0.75.$	$\sigma_1^2=0.5$, $\sigma_2^2=1.0.$	$\sigma_1^2=0.5$, $\sigma_2^2=2.0.$	$\sigma_1^2=0.75$, $\sigma_2^2=0.5.$	$\sigma_1^2=0.75$, $\sigma_2^2=0.75.$	$\sigma_1^2=0.75$, $\sigma_2^2=1.0.$	$\sigma_1^2=0.75$, $\sigma_2^2=2.0.$
Response (Y_1)	1	12.5000	11.9975	11.7561	12.1388	12.8741	13.2413	13.6406	12.1076	13.3516
	2	14.2857	13.4588	14.5418	14.4518	14.5286	13.5488	14.4832	15.4977	15.8118
	3	18.1818	18.9352	18.1558	17.7591	18.7179	18.8849	16.9437	17.7817	19.4180
	4	20.9302	20.4491	20.0622	20.9449	20.4410	21.5367	20.8006	20.9586	21.7198
	5	16.6667	15.4464	16.4721	16.9633	17.1476	17.3248	16.2294	17.3584	16.9497
	6	10.2740	10.8490	10.1605	11.1162	9.5156	8.7905	8.7765	11.0506	10.3343
	7	22.7273	23.7468	21.9611	23.2726	23.3635	24.0584	22.3657	22.8467	21.4284
	8	13.6364	14.1117	12.2545	11.7666	12.1351	12.2422	13.1038	12.2341	13.2742
	9	18.4211	18.5191	17.7779	18.6229	18.6224	19.3819	19.0453	16.9951	18.4028
	10	26.4206	25.8128	26.4167	27.0048	25.9021	25.4596	26.7145	26.7919	26.6184
	11	42.9078	42.3759	41.6891	42.9021	42.3609	43.2416	43.6724	42.2693	42.0347
	12	22.3372	23.2066	23.2303	22.9442	22.4445	23.1731	22.5833	22.8264	21.7616
	13	27.9365	28.7502	27.5119	28.4844	27.6983	28.6452	27.8105	26.7378	28.4199
	14	1.6129	1.1830	0.1535	2.5363	2.2993	1.6450	1.5353	2.0115	0.5836
	15	13.4146	13.9847	13.4931	14.2854	13.3388	12.6127	13.6651	13.9597	12.7430
	16	4.1985	4.3520	5.2504	4.8763	4.9151	4.1015	5.2072	4.5274	4.4332
	17	42.9078	42.6437	42.9453	41.7379	42.5717	42.2124	43.6056	42.0302	44.2371
	18	22.3372	21.7488	22.4517	21.6368	22.3859	20.8952	21.1631	22.0364	21.4259
	19	27.9365	28.1394	27.9175	28.4210	28.2184	27.1559	28.0412	28.3192	28.4783
Response (Y_2)	1	1.2500	-0.0362	1.4003	0.2751	2.8287	1.6660	1.0576	-0.3402	0.1220
	2	1.4286	0.3163	2.1926	0.8218	2.3060	1.8204	1.9237	0.7272	0.9852
	3	1.8182	3.2435	1.9761	2.5050	1.4113	1.5246	1.5583	0.7406	0.9665
	4	2.0930	2.0421	2.7471	2.1131	0.1530	2.1367	3.0753	3.0952	3.8736
	5	1.6667	3.5256	2.1066	2.7305	0.6967	1.9901	1.5113	3.3961	2.8807
	6	1.0274	0.8553	1.1416	-0.3137	1.4965	1.1681	-0.2431	1.7364	-1.9499
	7	2.2727	2.3953	2.5153	2.7522	0.8617	2.4548	3.4811	1.5248	1.7623
	8	1.3636	2.0164	0.5125	-0.2703	1.7758	2.8349	1.7454	1.5925	2.1465
	9	1.8421	1.7158	1.0245	0.3994	3.4078	0.2319	2.3317	1.6186	-0.3590
	10	2.6421	2.2732	2.6308	2.9358	2.9885	2.8818	2.0414	1.7888	2.3498
	11	4.2908	5.3033	4.5977	4.1504	4.5241	4.4958	5.0129	4.6364	3.6889
	12	2.2337	1.6185	1.4589	1.1034	2.8082	2.7020	0.2961	2.3435	2.9320
	13	2.7937	3.3647	3.4970	2.5011	4.5133	2.3829	3.7442	1.6606	1.5620
	14	0.1613	-0.1998	0.2562	-0.4212	2.2097	0.7890	0.1599	-0.5218	0.2742
	15	1.3415	1.8672	3.7071	0.4451	-0.1083	1.4630	-0.0568	1.0636	0.6038
	16	0.4198	1.0194	0.7759	0.6684	0.7104	1.0201	-0.6443	1.0746	-1.5797
	17	4.2908	3.7040	3.1590	2.8011	5.1236	4.9723	4.4704	3.0424	3.7473
	18	2.2337	2.6106	2.5661	2.5472	1.8603	3.1685	2.4251	1.6362	1.5861
	19	2.7937	3.5240	3.2262	0.7686	6.3226	2.7482	1.9224	2.3118	2.3814

	S.No.	σ_1^2=1.0, σ_2^2=0.5.	σ_1^2=1.0, σ_2^2=0.75.	σ_1^2=1.0, σ_2^2=1.0.	σ_1^2=1.0, σ_2^2=2.0.
Response (Y_1)	1	15.0807	12.8273	13.0551	13.1436
	2	12.9795	14.7614	12.3281	15.3052
	3	19.2053	18.5806	17.4214	19.1163
	4	21.7080	20.8575	18.4863	22.1588
	5	15.8328	17.9815	16.0073	16.4172
	6	9.6873	11.2523	10.1592	9.5664
	7	22.7929	24.4494	23.0274	22.1334
	8	13.6240	13.2241	13.0524	13.3741
	9	18.3441	18.9862	15.3473	19.6639
	10	24.8620	27.1605	27.9717	24.8717
	11	44.6104	43.1279	42.5004	42.5210
	12	21.8682	23.6500	23.7653	22.6123
	13	28.0311	28.5657	26.5833	28.7627
	14	1.9000	0.5049	2.5169	0.6336
	15	14.3340	12.9676	13.9563	13.3103
	16	4.7086	3.4725	3.7335	4.3263
	17	43.1532	43.2618	45.3382	42.9703
	18	20.9366	21.8304	24.3576	22.7088
	19	28.9061	25.8328	28.7338	27.8325
Response (Y_2)	1	2.3769	0.6744	1.2810	0.2646
	2	0.4118	2.6844	1.9693	0.8815
	3	0.7333	1.5326	2.5021	1.8410
	4	2.0402	4.4329	1.5030	4.0294
	5	1.7243	0.2508	1.4056	2.2562
	6	0.4311	0.5629	2.5445	1.1246
	7	1.8735	2.7466	3.2800	2.6888
	8	1.3437	2.6563	1.6671	2.0320
	9	0.9627	1.4479	1.0250	4.3633
	10	3.1604	2.2122	2.1509	3.0106
	11	4.3329	5.3605	5.1582	6.4273
	12	2.3392	2.2692	2.5945	2.6820
	13	3.9221	3.4418	2.7133	3.9299
	14	-0.3883	1.2272	0.9106	0.9828
	15	2.4377	1.6055	-0.4505	3.8568
	16	1.1658	1.3532	1.6331	-0.9068
	17	4.1729	4.9534	4.2303	2.9944
	18	2.4561	1.0909	1.8412	2.7655
	19	3.7974	2.3527	3.4032	4.0803

APPENDIX II

The subroutine used for the minimization of objective function by GNLM method is given in this section. The code is written in MATLAB.

```
function [beta,r,J] = gsnw(X,y,model,beta0)
global s
%function for Guass Newton-Levenberg Marquardt
Method
if (nargin<4)
error('GNLM requires four arguments.'); end
if min(size(y)) ~= 1
   error('Requires a vector second input
argument.');
end
y = y(:);
a = size(y,1)/2;
s1 = sqrt(s(1));
s2 = sqrt(s(2));
b1 = repmat(s1,a,1);
b2 = repmat(s2,a,1);
b = [b1;b2];
if size(X,1) == 1 % turn a row vector into a column
vector.
   X = X(:);
end
wasnan = (isnan(y) | any(isnan(X),2));
if (any(wasnan))
   y(wasnan) = [ ];
   X(wasnan,:) = [ ];
end
n = length(y);
p = length(beta0);
beta0 = beta0(:);
J = zeros(n,p);
beta = beta0;
betanew = beta + 1;
maxiter = 100;
iter = 0;
betatol = 1.0E-4;
rtol = 1.0E-4;
sse = 1;
sseold = sse;
seps = sqrt(eps);
zbeta = zeros(size(beta));
s10 = sqrt(10);
eyep = eye(p);
zerosp = zeros(p,1);
```

```
while (norm((betanew-beta)./(beta+seps)) > betatol |
abs(sseold-sse) > rtol) & iter < maxiter
   if iter > 0,
      beta = betanew;
   end
   iter = iter + 1;
   yfit = feval(model,beta,X);
   r = y - yfit;
   r = r.*b;
   sseold = r'*r;
   for k = 1:p,
      delta = zbeta;
      if (beta(k) == 0)
         nb = sqrt(norm(beta));
         delta(k) = seps * (nb + (nb==0));
      else
         delta(k) = seps*beta(k);
      end
      yplus = feval(model,beta+delta,X);
      J(:,k) = (yplus - yfit)/delta(k);
   end
   Jplus = [J;(1.0E-2)*eyep];
   rplus = [r;zerosp];

   % Levenberg-Marquardt type adjustment
   % Gauss-Newton step -> J\r
   % LM step -> inv(J'*J+constant*eye(p))*J'*r
   step = Jplus\rplus;
   betanew = beta + step;
   yfitnew = feval(model,betanew,X);
   rnew = y - yfitnew;
   rnew = rnew.*b;
   sse = rnew'*rnew;
   iter1 = 0;
   while sse > sseold & iter1 < 20
      step = step/s10;
      betanew = beta + step;
      yfitnew = feval(model,betanew,X);
      rnew = y - yfitnew;
      rnew = rnew.*b;
      sse = rnew'*rnew;
      iter1 = iter1 + 1;
   end
end
if iter == maxiter
   disp('GNLM did NOT converge. Returning results
from last iteration.');
end
```

APPENDIX III

The program used for the simulation of Single response system is given in this section. This program includes all the calculations of average estimates, mean squared errors, confidence intervals and coverage probabilities by asymptotic theory and bootstrapping methods. The code is written in MATLAB.

```
clc;
global k dt
data;
y = dt(:,3);
c = dt(:,1:2);
fp = fopen('sim5.txt','a');
fp1 = fopen('data5.txt','a');
sg = 1;
%while sg<0.06
x0 = [400 5000];
i = 0;
pa(1:4) = 0;
pb(1:4) = 0;
while i <1000
 i = i+1
 er = randn(12,1);
 yt = y+er*sg;
 %Bootstrapping Perameters
 d = [yt c];
 bst = bootstrp(100,'m2fit',d);
 mbt(i,:) = mean(bst);
 %Bootstrap Confidence Limits
 sbt = sort(bst);
 lb = (sbt(2,:)+sbt(3,:))/2;
 ub = (sbt(97,:)+sbt(98,:))/2;
 cb = [lb' ub'];
 bt(i,:) = [cb(1,:) cb(2,:)];
 k = 1;
 if bt(i,2*k-1)>400 | bt(i,2*k)<400
     pb(k) = pb(k)+1;
 end
 k = k+1;
 if bt(i,2*k-1)>5000 | bt(i,2*k)<5000
     pb(k) = pb(k)+1;
 end
 %Calculation of Perameters
 yt = yt(:);
 [x er J] = nlinfit(c,yt,@m2fun,x0);
 est(i,:) = x';
 fprintf(fp1,'%8.4f ', est(i,:));
```

```
%Assymptotic Confidence Limits
e = er'*er;
a = diag(inv(J'*J));
ns = size(yt,1);
b = 1.96.*sqrt(a.*e/(ns-3));
cc = [x-b x+b];
at(i,:) = [cc(1,:) cc(2,:)];
fprintf(fp1,'%8.4f ', at(i,:));
fprintf(fp1,'\n');
k = 1;
if at(i,2*k-1)>400 | at(i,2*k)<400
    pa(k) = pa(k)+1;
end
k = k+1;
if at(i,2*k-1)>5000 | at(i,2*k)<5000
    pa(k) = pa(k)+1;
end
end
mt = mean(est);
fprintf(fp,'\nMean Pearameters\n');
fprintf(fp,'%12.4f ',mt);
mse = diag(cov(est));
fprintf(fp,'\nMean Squared Errors of Perameters\n');
fprintf(fp,'%12.4f ',mse);
j = 0;
while j < 2
j = j+1;
bi(:,j) = bt(:,2*j)-bt(:,2*j-1);
ai(:,j) = at(:,2*j)-at(:,2*j-1);
end
mbi = mean(bi);
mai = mean(ai);
fprintf(fp,'\nMean length of ACI\n');
fprintf(fp,'%8.4f ',mai);
fprintf(fp,'\nMean length of BtCI \n');
fprintf(fp,'%8.4f ',mbi);
pa = 1000-pa;
fprintf(fp,'\nCoverage Probabilities of ACI \n');
fprintf(fp,'%4d ',pa);
pb = 1000-pb;
fprintf(fp,'\nCoverage Probabilities of BtCI \n');
fprintf(fp,'%4d ',pb);
  k = 0;
  s = 100;
  qqpt(est(:,1));
  str = ['qqp',int2str(s),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  histfit(est(:,1),30)
```

```
    str = ['hst',int2str(s),'K',int2str(k)];
    saveas(gcf, str , 'fig')
    qqpt(est(:,2));
    str = ['qqp',int2str(s),'K',int2str(k)];
    saveas(gcf, str , 'fig')
    histfit(est(:,2),30)
    str = ['hst',int2str(s),'K',int2str(k)];
    saveas(gcf, str , 'fig')
fprintf(fp1,'\n\n\n');
%end
fclose(fp);
fclose(fp1);
disp('PROGRAMME END.');

function x=m2fit(dt)
y = dt(:,1);
c = [dt(:,2) dt(:,3)];
x0 = [400 5000];
x = nlinfit(c,y,@m2fun,x0);
x =x';

function yhat=m2fun(s,c)
n = ones(12,1);
d = n+s(1).*c(:,1).*exp(-s(2)./c(:,2));
yhat = n./d;
```

APPENDIX IV

The program used for the simulation of Multi response system is given in this section. This program includes all the calculations of average estimates, mean squared errors, confidence intervals and coverage probabilities by asymptotic theory and bootstrapping methods. The code is written in MATLAB.

```
clc;
global s k dt
data;
y = dt(:,3:4);
c = dt(:,1:2);
ii = 0.75;
fp = fopen('sim2.txt','a');
fp1 = fopen('data3.txt','a');
jj = 1.0;
%while jj < 2.1
if jj == 1.25
   jj = 2.0;
end
s = [1/ii 1/jj];
mu = [0 0];
cv = [ii 0;0 jj];
fprintf(fp,'\n\nFor variences: %4.2f %4.2f',ii,jj);
n = 10;
while n<11
if n==20
   n=19;
end
fprintf(fp,'\n\nResults for sample size: %d\n',n);
c1 = cn(1:n,:);
y = yn(1:n,:);
c = repmat(c1,2,1);
x0 = [1 0.1 1 0.1];
i = 0;
pa(1:4) = 0;
pb(1:4) = 0;
while i <1000
 i = i+1;
 [n i]
 er = mvrnd(mu,cv,n);
 yt = y+er;
 %Bootstrapping Perameters
 d = [c1 yt];
 bst = bootstrp(100,'crfit',d);
 mbt(i,:) = mean(bst);
 %Bootstrap Confidence Limits
```

```
sbt = sort(bst);
lb = (sbt(2,:)+sbt(3,:))/2;
ub = (sbt(97,:)+sbt(98,:))/2;
cb = [lb' ub'];
bt(i,:) = [cb(1,:) cb(2,:) cb(3,:) cb(4,:)];
k = 1;
while k<5
if bt(i,2*k-1)>0.1 | bt(i,2*k)<0.1
    pb(k) = pb(k)+1;
end
k = k+1;
if bt(i,2*k-1)>0.01 | bt(i,2*k)<0.01
    pb(k) = pb(k)+1;
end
k = k+1;
end
%Calculation of Perameters
yt = yt(:);
[x er J] = gsnw(c,yt,@p1fun,x0);
est(i,:) = x';
fprintf(fp1,'%8.4f', est(i,:));
fprintf(fp1,'\n');
%Assymptotic Confidence Limits
y1 = p1fun(x,c);
e1 = y1-yt;
e = e1'*e1;
a = diag(inv(J'*J));
ns = size(y1,1);
b = 1.96.*sqrt(a.*e/(ns-5));
cc = [x-b x+b];
at(i,:) = [cc(1,:) cc(2,:) cc(3,:) cc(4,:)];
k = 1;
while k<5
if at(i,2*k-1)>0.1 | at(i,2*k)<0.1
    pa(k) = pa(k)+1;
end
k = k+1;
if at(i,2*k-1)>0.01 | at(i,2*k)<0.01
    pa(k) = pa(k)+1;
end
k = k+1;
end
end
mt = mean(est);
fprintf(fp,'\nMean Pearameters\n');
fprintf(fp,'%8.4f',mt);
mse = diag(cov(est))*1e4;
fprintf(fp,'\nMean Squared Errors of Perameters\n');
```

```
fprintf(fp,'%12.4f',mse);
j = 0;
while j < 4
j = j+1;
bi(:,j) = bt(:,2*j)-bt(:,2*j-1);
ai(:,j) = at(:,2*j)-at(:,2*j-1);
end
mbi = mean(bi);
mai = mean(ai);
fprintf(fp,'\nMean length of ACI\n');
fprintf(fp,'%8.4f',mai);
fprintf(fp,'\nMean length of BtCI \n');
fprintf(fp,'%8.4f',mbi);
pa = 1000-pa;
fprintf(fp,'\nCoverage Probabilities of ACI \n');
fprintf(fp,'%4d',pa);
pb = 1000-pb;
fprintf(fp,'\nCoverage Probabilities of BtCI \n');
fprintf(fp,'%4d',pb);
k = 0;
  qqpt(est(:,1));
  str = ['qqp',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  histfit(est(:,1),30)
  str = ['hst',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  qqpt(est(:,2));
  str = ['qqp',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  histfit(est(:,2),30)
  str = ['hst',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  qqpt(est(:,3));
  str = ['qqp',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  histfit(est(:,3),30)
  str = ['hst',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  qqpt(est(:,4));
  str = ['qqp',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
  histfit(est(:,4),30)
  str = ['hst',int2str(n),'K',int2str(k)];
  saveas(gcf, str , 'fig')
n = n+5;
fprintf(fp1,'\n\n\n');
end
jj = jj+0.25;
```

```
%end
fclose(fp);
fclose(fp1);
disp('PROGRAMME END.');

function x=crfit(d)
global fp
dt = d;
y = [dt(:,3) dt(:,4)];
y = y(:);
c = [dt(:,1) dt(:,2); dt(:,1) dt(:,2)];
x0 = [0.1 0.01 0.1 0.01];
x = gsnw(c,y,@p1fun,x0);
x = x';

function [yhat,d]=p1fun(s,c1)
a = size(c1,1)/2;
c = c1(1:a,:);
x1 = c(:,1);
x2 = c(:,2);
o = ones(size(x1));
d = o+s(3).*x1+s(4).*x2;
y1 = (s(1).*x1.*x2)./d;
y2 = (s(2).*x1.*x2)./d;
yhat = [y1;y2];
```

APPENDIX V

The subroutine used for the generation of multivariate random number is given here. The code is written in MATLAB.

```
function X = mvrnd(mu,covm,n)
% MVRND Generate multivariate normal random
variables.
%
%   R = MVRND(MU,COVM,N) Generates a sample of size
N
%   random variables from the multivariate normal
%   distribution. MU is the d-dimensional mean as a
%    column vector. COVM is the d x d covariance
matrix.

if det(covm) <=0
    % Then it is not a valid covariance matrix.
    error('The covariance matrix must be positive
definite')
end

mu = mu(:); % Just in case it is not a column
vector.
d = length(mu);
% get cholesky factorization of covariance
R = chol(covm);
% generate the standard normal random variables
Z = randn(n,d);
X = Z*R + ones(n,1)*mu';
```

APPENDIX VI

The program used for the cross validation procedure for single response system is given here. The code is written in MATLAB.

```
clear all
clc
data
dt1 = [y1 c1];
cv = 0;
n=length(y1);
for i = 1:n
 ii = [(1:(i-1)) ((i+1):n)];
 dt = dt1(ii,:);
 y = dt(:,1);
 c = [dt(:,2) dt(:,3)];
 x0 = [400 5000];
 x = nlinfit(c,y,@m3fun,x0);
 yhat = exp(-x(1).*c1(i,1).*exp(-x(2)./c1(i,2)));
 %yhat = 1/(1+x(1).*c1(i,1).*exp(-x(2)./c1(i,2)));
 %yhat = (1+2*x(1).*c1(i,1).*exp(-x(2)./c1(i,2)))^
(-0.5);
 %yhat = (1+3*x(1).*c1(i,1).*exp(-x(2)./c1(i,2)))^
(-1/3);
 cv = cv + (yhat-y1(i))^2;
 i
end
cv

function yhat=m3fun(s,c)
%model1
yhat = exp(-s(1).*c(:,1).*exp(-s(2)./c(:,2)));
%model3 & 4
%n = ones(size(c,1),1);
%d = n+2.*s(1).*c(:,1).*exp(-s(2)./c(:,2));
%yhat = d.^(-1/3);
```

APPENDIX VII

The program used for the cross validation procedure for multi response system is given here. The code is written in MATLAB.

```
global s
clc
data;
ii = 0.5;
jj = 2;
n = 19;
s = [1/ii 1/jj];
x0 = [1 0.1 1 0.1];
yt = yn;
dt1 = [yt cn];
cv = 0;
for i = 1:n
     ii = [(1:(i-1)) ((i+1):n)];
     dt = dt1(ii,:);
     y = [dt(:,1) dt(:,2)];
     c = [dt(:,3) dt(:,4)];
     c = repmat(c,2,1);
     y = y(:);
     x = gsnw(c,y,@mrf1,x0);
     d1 = (1 + x(3)*cn(i,1)+x(4)*cn(i,2));
     d2 = (1 + x(3)*cn(i,1)+x(4)*cn(i,2));
     yh(1) = (x(1)*cn(i,1)*cn(i,2))/d1;
     yh(2) = (x(2)*cn(i,1)*cn(i,2))/d2;
     cv = cv + sum((yh-yn(i,:)).^2)
     i
 end
cv

function yhat=mrf1(s,c1)
a = size(c1,1)/2;
c = c1(1:a,:);
x1 = c(:,1);
x2 = c(:,2);
o = ones(size(x1));
d = o+s(3).*x1+s(4).*x2;
y1 = (s(1).*x1.*x2)./d;
y2 = (s(2).*x1.*x2)./d;
yhat = [y1;y2];
```